

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 12, 2006, 11:45:28 ; Search time 74.4018 Seconds  
(without alignments)  
822.262 Million cell updates/sec

Title: US-09-581-651D-41  
Perfect score: 54  
Sequence: 1 VSIPRNLGY 10

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 9312410 seqs, 2039259788 residues

Total number of hits satisfying chosen parameters: 18624820

Minimum DB seq length: 0  
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/abs/ABSWEB.spool/US09581651/runat\_10052006\_181225\_634/app\_query.fasta.1  
-DB=Published Applications NA New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62  
-TRANS-human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -HOST=abes06p  
-USER=US09581651 @CGN 1.1.1490 @runat\_10052006\_181225\_634 -NCPUs=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DISPLOCK=100 -LONGLOG -DEV TIMOUT=120  
-WARN TIMOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA New:  
1: /SIDSS/ptodata/2/pubpna/US08 NEW PUB.seq1.\*  
2: /SIDSS/ptodata/2/pubpna/US06 NEW PUB.seq.\*  
3: /SIDSS/ptodata/2/pubpna/US07 NEW PUB.seq.\*  
4: /SIDSS/ptodata/2/pubpna/US08 NEW PUB.seq.\*  
5: /SIDSS/ptodata/2/pubpna/PCT NEW PUB.seq.\*  
6: /SIDSS/ptodata/2/pubpna/US09 NEW PUB.seq.\*  
7: /SIDSS/ptodata/2/pubpna/US09 NEW PUB.seq1.\*  
8: /SIDSS/ptodata/2/pubpna/US09 NEW PUB.seq2.\*  
9: /SIDSS/ptodata/2/pubpna/US10 NEW PUB.seq.\*  
10: /SIDSS/ptodata/2/pubpna/US10 NEW PUB.seq1.\*  
11: /SIDSS/ptodata/2/pubpna/US10 NEW PUB.seq2.\*  
12: /SIDSS/ptodata/2/pubpna/US10 NEW PUB.seq3.\*  
13: /SIDSS/ptodata/2/pubpna/US11 NEW PUB.seq.\*  
14: /SIDSS/ptodata/2/pubpna/US11 NEW PUB.seq1.\*  
15: /SIDSS/ptodata/2/pubpna/US11 NEW PUB.seq2.\*  
16: /SIDSS/ptodata/2/pubpna/US11 NEW PUB.seq3.\*  
17: /SIDSS/ptodata/2/pubpna/US11 NEW PUB.seq4.\*  
18: /SIDSS/ptodata/2/pubpna/US11 NEW PUB.seq5.\*  
19: /SIDSS/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result \*

Query

No.	Score	Match	Length	DB	ID	Description
1	54	100.0	2402	17	US-11-193-561-26	Sequence 26, Appl
2	54	100.0	2402	17	US-11-193-771-26	Sequence 26, Appl
3	54	100.0	2402	17	US-11-193-789-26	Sequence 26, Appl
4	54	100.0	2402	17	US-11-193-806-26	Sequence 26, Appl
5	54	100.0	2402	17	US-11-193-857-26	Sequence 114, App
6	54	100.0	2443	10	US-10-995-561-114	Sequence 114, App
7	54	100.0	2488	10	US-10-995-561-105	Sequence 105, App
8	54	100.0	87672	10	US-10-995-561-13237	Sequence 13237, A
9	54	100.0	143947	17	US-11-193-561-37	Sequence 37, Appl
10	54	100.0	143947	17	US-11-193-771-37	Sequence 37, Appl
11	54	100.0	143947	17	US-11-193-789-37	Sequence 37, Appl
12	54	100.0	143947	17	US-11-193-806-37	Sequence 37, Appl
13	54	100.0	143947	17	US-11-193-857-37	Sequence 37, Appl
14	45	83.3	79084	11	US-10-330-773-540	Sequence 540, App
15	43	79.6	927	17	US-11-102-026A-88	Sequence 88, Appl
16	42	77.8	25	17	US-11-121-849-485840	Sequence 485840, A
17	42	77.8	613	7	US-09-925-065A-834813	Sequence 834813, A
18	42	77.8	616	7	US-09-925-065A-829110	Sequence 829110, A
19	42	77.8	617	7	US-09-925-065A-840083	Sequence 840083, A
20	42	77.8	1498	10	US-10-750-185-53394	Sequence 53394, A
21	42	77.8	1498	10	US-10-750-623-53394	Sequence 53394, A
22	41	75.9	654	7	US-09-925-065A-835700	Sequence 835700, A
23	40	74.1	527	7	US-09-925-065A-824091	Sequence 824091, A
24	40	74.1	533	7	US-09-925-065A-831462	Sequence 831462, A
25	40	74.1	578	7	US-09-925-065A-365691	Sequence 365691, A
26	40	74.1	586	12	US-10-301-480-436207	Sequence 436207, A
27	40	74.1	586	12	US-10-301-480-1049616	Sequence 1049616, A
28	40	74.1	1480	18	US-11-072-512-128	Sequence 128, App
29	40	74.1	1626	11	US-10-301-480-97216	Sequence 97216, A
30	40	74.1	1626	12	US-10-301-480-710625	Sequence 710625, A
31	40	74.1	1626	12	US-10-301-480-553082	Sequence 553082, A
32	40	74.1	2674	7	US-09-925-065A-100	Sequence 100, App
33	40	74.1	124972	17	US-11-121-086-151	Sequence 151, App
34	40	74.1	128361	9	US-10-505-928-151	Sequence 151, App
35	40	74.1	1694969	13	US-10-506-454-1690	Sequence 1690, Ap
36	39	72.2	570	7	US-09-925-065A-444743	Sequence 444743, A
37	39	72.2	575	12	US-10-301-480-504951	Sequence 504951, A
38	39	72.2	575	12	US-10-301-480-1118360	Sequence 1118360, A
39	39	72.2	601	11	US-10-301-480-52837	Sequence 52837, A
40	39	72.2	601	12	US-10-301-480-666246	Sequence 666246, A
41	39	72.2	604	7	US-09-925-065A-610379	Sequence 610379, A
42	39	72.2	607	7	US-09-925-065A-272220	Sequence 272220, A
43	39	72.2	607	7	US-09-925-065A-272221	Sequence 272221, A
44	39	72.2	607	12	US-10-301-480-349526	Sequence 349526, A
45	39	72.2	607	12	US-10-301-480-349527	Sequence 349527, A

ALIGNMENTS

RESULT 1

US-11-193-561-26  
; Sequence 26, Application US/11193561  
; Publication No. US20060024757A1  
; GENERAL INFORMATION:  
; APPLICANT: Husa, Robert  
; APPLICANT: Shorter, Simon  
; TITLE OF INVENTION: Detection of Oncofetal Fibronectin for Selection of Concepti  
; FILE REFERENCE: 17101-080001/831  
; CURRENT APPLICATION NUMBER: US/11/193,561  
; PRIOR FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: 60/592,823  
; PRIOR FILING DATE: 2004-07-30  
; PRIOR APPLICATION NUMBER: 60/592,803  
; PRIOR FILING DATE: 2004-07-30  
; PRIOR APPLICATION NUMBER: 60/592,825  
; PRIOR FILING DATE: 2004-07-30  
; PRIOR APPLICATION NUMBER: 60/592,804  
; PRIOR FILING DATE: 2004-07-30  
; PRIOR APPLICATION NUMBER: 60/592,824  
; PRIOR FILING DATE: 2004-07-30  
; NUMBER OF SEQ ID NOS: 39

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 2402
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: Oncofetal Fibronectin
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank NM 051
; DATABASE ENTRY DATE: 2005-08-10
US-11-193-561-26

```

Alignment Scores:		
Pred. No.:	8-29	2402
Score:	54.00	Matches: 10
Percent Similarity:	100.0%	Conservative: 0
Best Local Similarity:	100.0%	Mismatches: 0
Query Match:	100.0%	Indels: 0
DB:	17	Gaps: 0

US-09-581-651D-41 (1-10) x US-11-193-561-26 (1-2402)

Qy 1 ValSerIleProProArgAsnLeuGlyTyr 10  
2208 GTGAGTATCCCAACCAGAAACCTTGGATAC 2237

## RESULT 2

US-11-193-771-26  
; Sequence 26, Application US/11193771  
; Publication No. US20060024722A1

Alignment Scores:	
Pred. No.:	8.29
Score:	54.00
Percent Similarity:	100.0%
Best Local Similarity:	100.0%
Query Match:	100.0%
DB:	17
Length:	2402
Matches:	10
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-581-651D-41 (1-10) x US-11-193-771-26 (1-2402)

Qy 1 ValSerIleProProArgAsnLeuGlyTyr 10

db 2208 GTGAGTATCCCACCCAGAAACCTTGGATAC 2237

### RESULT 3

US-11-193-789-26  
; Sequence 26, Application US/11193789  
; Publication No. US20060024723A1

Alignment Scores:		
Pred. No.:	8.29	Length:
Score:	54.00	Matches:
Percent Similarity:	100.0%	Conservative:
Best Local Similarity:	100.0%	Mismatches:
Query Match:	100.0%	Indels:
DB:	17	Gaps:
		2402

US-09-581-651D-41 (1-10) X US-11-193-789-26 (1-2402)

Qy	1	valSerIleProProArgAsnLeuClyTyr	10
Db	2208	GTGAGTATCCCAACCCAGAACCTTGATAC	2229

## RESULT 4

US-11-193-806-26 ; Sequence 26, Application US/11193806  
: Publication No. US20060024724A1

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 12, 2006, 11:24:42 ; Search time 23.773 Seconds  
(without alignments)  
5217.711 Million cell updates/sec

Title: US-09-581-651D-41

Perfect score: 54

Sequence: 1 VSIPPNLGY 10

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/abs/ABSSWEB.spool/US09581651/runat.10052006.181222.525/app.query.fasta.1  
-DB=Published Applications NA.Main -OPMT=fastap -SUFFIX=p2n.rnpbm  
-MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1  
-MATRIX=biosum62 -TRANS=human40.cdi -LIST=45 -MODE=LOCAL -OUTFMT=ptc -NORM=ext  
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext  
-HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abes03p  
-USR=US09581651 @CGN 1.1 2064 @runat.10052006.181222.525 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DISBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOG=6 -DELEXT=7

Database : Published Applications NA.Main:

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	100.0	471	7	US-10-242-535A-42467 Sequence 42467, A
2	54	100.0	471	7	US-10-085-783A-42467 Sequence 42467, A
3	54	100.0	600	9	US-10-956-157-9523 Sequence 9523, Ap
4	54	100.0	2127	6	US-10-210-120-49 Sequence 49, Appl
5	54	100.0	2127	9	US-10-956-157-4288 Sequence 4288, Ap
6	54	100.0	2127	9	US-10-909-035-49 Sequence 49, Appl
7	54	100.0	2443	7	US-10-741-601-70 Sequence 70, Appl

8	54	100.0	2443	8	US-10-741-600-238 Sequence 238, App
9	54	100.0	2488	7	US-10-741-601-75 Sequence 75, Appl
10	54	100.0	2488	8	US-10-741-600-244 Sequence 244, App
11	54	100.0	4295	6	US-10-144-194A-51 Sequence 51, Appl
12	54	100.0	4295	8	US-10-491-566-51 Sequence 51, Appl
13	54	100.0	87467	7	US-10-741-601-5634 Sequence 5634, Ap
14	54	100.0	87467	8	US-10-741-601-17624 Sequence 174961, A
c 15	43	79.6	3186778	5	US-10-027-632-174961 Sequence 174961, A
c 16	43	79.6	3186778	6	US-10-027-632-174961 Sequence 174961, A
c 17	42	77.8	418	8	US-10-425-115-131037 Sequence 131037, A
c 18	42	77.8	418	8	US-10-425-115-131037 Sequence 131037, A
c 19	42	77.8	616	4	US-09-925-065A-834813 Sequence 834813, A
c 20	42	77.8	616	4	US-09-925-065A-834813 Sequence 834813, A
c 21	42	77.8	617	4	US-09-925-065A-829110 Sequence 829110, A
c 22	42	77.8	2234	3	US-09-813-153-52 Sequence 840083, A
c 23	42	77.8	2234	4	US-09-949-925-52 Sequence 52, Appl
c 24	42	77.8	8628	3	US-09-960-706-952 Sequence 952, App
c 25	42	77.8	8628	3	US-09-960-706-952 Sequence 952, App
c 26	42	77.8	8955	10	US-11-097-143-33433 Sequence 622, App
c 27	41	75.9	3330	10	US-11-097-143-33433 Sequence 33433, A
c 28	41	75.9	654	4	US-10-425-115-36206 Sequence 835700, A
c 29	41	75.9	1110	10	US-11-097-143-15701 Sequence 15701, A
c 30	41	75.9	2285	10	US-11-097-143-28162 Sequence 28162, A
c 31	41	75.9	3330	10	US-11-097-143-15700 Sequence 15700, A
c 32	41	75.9	228139	5	US-10-087-192-232 Sequence 232, App
c 33	40	74.1	492	5	US-10-027-632-281008 Sequence 281008, A
c 34	40	74.1	492	6	US-10-027-632-281008 Sequence 281008, A
c 35	40	74.1	527	4	US-09-925-065A-824091 Sequence 824091, A
c 36	40	74.1	527	4	US-09-925-065A-824091 Sequence 824091, A
c 37	40	74.1	578	4	US-09-925-065A-831462 Sequence 831462, A
c 38	40	74.1	578	4	US-09-925-065A-831462 Sequence 831462, A
c 39	40	74.1	662	3	US-09-969-034-815 Sequence 79249, A
c 40	40	74.1	662	3	US-09-969-034-815 Sequence 79249, A
c 41	40	74.1	728	6	US-10-133-013-21 Sequence 21, Appl
c 42	40	74.1	728	6	US-10-133-013-21 Sequence 21, Appl
c 43	40	74.1	941	10	US-10-027-632-99730 Sequence 99730, A
c 44	40	74.1	941	10	US-10-027-632-99730 Sequence 99730, A
c 45	40	74.1	1480	6	US-11-097-143-11702 Sequence 11702, A
			1542	9	US-10-104-047-128 Sequence 128, App
					Sequence 5, Appl

#### ALIGNMENTS

RESULT 1  
US-10-242-535A-42467  
; Sequence 42467, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2005  
; CURRENT APPLICATION NUMBER: US/10/242,535A  
; PRIOR FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 10/085,783  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 42467  
; LENGTH: 471  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (7)..(7)  
; OTHER INFORMATION: n is a, c, g, or t  
US-10-242-535A-42467

Alignment Scores:  
Pred. No.: 1-29 Length: 471  
Score: 54.00 Matches: 10  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 7 Gaps: 0

US-09-581-651D-41 (1-10) x US-10-242-535A-42467 (1-471)

Qy 1 ValSerIleProProArgAsnLeuGlyTyr 10  
Db 393 GTGAGTATCCACCCAGAAACCTTGGATAC 422

RESULT 2  
US-10-085-783A-42467  
; Sequence 42467, Application US/10085783A  
; Publication No. US20040037841A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2002  
; CURRENT APPLICATION NUMBER: US/10/085,783A  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 59994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 42467  
; LENGTH: 471  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (7)..(7)  
; OTHER INFORMATION: n is a, c, g, or t

US-10-085-783A-42467

Alignment Scores:  
Pred. No.: 1-29 Length: 471  
Score: 54.00 Matches: 10  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 7 Gaps: 0

US-09-581-651D-41 (1-10) x US-10-085-783A-42467 (1-471)

Qy 1 ValSerIleProProArgAsnLeuGlyTyr 10  
Db 393 GTGAGTATCCACCCAGAAACCTTGGATAC 422

RESULT 3  
US-10-956-157-9523  
; Sequence 9523, Application US/10956157  
; Publication No. US20050118625A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Mounts, William  
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH  
; FILE REFERENCE: 031896-043000 (AM 101081)  
; CURRENT APPLICATION NUMBER: US/10/956,157  
; CURRENT FILING DATE: 2004-10-04  
; NUMBER OF SEQ ID NOS: 319805  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 9523  
; LENGTH: 600

; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-956-157-9523

Alignment Scores:  
Pred. No.: 1-65 Length: 600  
Score: 54.00 Matches: 10  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 9 Gaps: 0

US-09-581-651D-41 (1-10) x US-10-956-157-9523 (1-600)

Qy 1 ValSerIleProProArgAsnLeuGlyTyr 10  
Db 426 GTGAGTATCCACCCAGAAACCTTGGATAC 455

RESULT 4  
US-10-210-120-49  
; Sequence 49, Application US/10210120  
; Publication No. US20030175736A1  
; GENERAL INFORMATION:  
; APPLICANT: Chinnaiyan, Arul M.  
; APPLICANT: Rubin, Mark A.  
; APPLICANT: Sreekumar, Arun  
; TITLE OF INVENTION: Expression Profile of Prostate Cancer  
; FILE REFERENCE: UM-07221  
; CURRENT APPLICATION NUMBER: US/10/210,120  
; CURRENT FILING DATE: 2002-08-01  
; PRIOR APPLICATION NUMBER: US 60/309,581  
; PRIOR FILING DATE: 2001-08-02  
; PRIOR APPLICATION NUMBER: US 60/334,468  
; PRIOR FILING DATE: 2001-11-15  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 49  
; LENGTH: 2127  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-10-210-120-49

Alignment Scores:  
Pred. No.: 5-95 Length: 2127  
Score: 54.00 Matches: 10  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 6 Gaps: 0

US-09-581-651D-41 (1-10) x US-10-210-120-49 (1-2127)

Qy 1 ValSerIleProProArgAsnLeuGlyTyr 10  
Db 1953 GTGAGTATCCACCCAGAAACCTTGGATAC 1982

RESULT 5  
US-10-956-157-4288  
; Sequence 4288, Application US/10956157  
; Publication No. US20050118625A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Mounts, William  
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH  
; FILE REFERENCE: 031896-043000 (AM 101081)  
; CURRENT APPLICATION NUMBER: US/10/956,157  
; CURRENT FILING DATE: 2004-10-04  
; NUMBER OF SEQ ID NOS: 319805  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4288  
; LENGTH: 2127  
; TYPE: DNA

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 12, 2006, 10:33:26 ; Search time 6.42638 Seconds  
(without alignments)  
4149.058 Million cell updates/sec

Title: US-09-581-651D-41  
Perfect score: 54  
Sequence: 1 VSIPPNLGY 10

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/abs/ABSSWEB.spool/US09581651/runat\_10052006\_181213\_253/app\_query.fasta\_1  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-HOST=abs804 -USER=US09581651 @CGN 1 1 237 @runat 10052006 181213 253 -NCPU=6  
-ICPU=3 -NO WMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.\*

- 1: /cgn2\_6/ptodata/1/ina/1 COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5 COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/H COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*
- 7: /cgn2\_6/ptodata/1/ina/PP COMB.seq.\*
- 8: /cgn2\_6/ptodata/1/ina/RE COMB.seq.\*
- 9: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	75.9	601	3	US-09-949-016-154321
2	41	75.9	601	3	US-09-949-016-154322
3	41	75.9	601	3	US-09-949-016-154323
4	41	75.9	212139	3	US-09-949-016-16065
C 5	40	74.1	426	3	US-09-621-976-16065
C 6	40	74.1	449	3	US-09-621-976-19038
C 7	40	74.1	1480	3	US-10-104-047-128
C 8	40	74.1	1930	3	US-09-919-172-3
C 9	40	74.1	1930	3	US-09-976-594-957

10	40	74.1	26086	3	US-09-949-016-15355	Sequence 15355, A
11	40	74.1	26238	3	US-09-949-016-12314	Sequence 12314, A
12	39	72.2	601	3	US-09-949-016-105318	Sequence 105318, A
C 13	39	72.2	3003	3	US-08-915-337-1	Sequence 1, Appli
C 14	39	72.2	19818	3	US-09-949-016-12198	Sequence 12198, A
C 15	39	72.2	19819	3	US-09-949-016-16987	Sequence 16987, A
16	39	72.2	41617	3	US-09-949-016-14356	Sequence 14356, A
17	39	72.2	41618	3	US-09-949-016-14681	Sequence 14681, A
C 18	39	72.2	58108	3	US-09-949-016-13383	Sequence 13383, A
C 19	39	72.2	231129	3	US-09-949-016-16110	Sequence 16110, A
C 20	39	72.2	266293	3	US-09-949-016-11934	Sequence 11934, A
21	38	70.4	502	3	US-09-270-767-6425	Sequence 6425, Ap
22	38	70.4	502	3	US-09-270-767-21707	Sequence 21707, A
23	38	70.4	601	3	US-09-949-016-27189	Sequence 27189, A
C 24	38	70.4	601	3	US-09-949-016-40844	Sequence 40844, A
C 25	38	70.4	601	3	US-09-949-016-40845	Sequence 40845, A
C 26	38	70.4	601	3	US-09-949-016-135102	Sequence 135102, A
C 27	38	70.4	601	3	US-09-949-016-182406	Sequence 182406, A
C 28	38	70.4	601	3	US-09-949-016-200981	Sequence 200981, A
29	38	70.4	607	3	US-09-270-767-10444	Sequence 10444, A
30	38	70.4	780	3	US-09-540-236-1835	Sequence 1835, Ap
31	38	70.4	807	3	US-09-252-991A-11725	Sequence 11725, A
32	38	70.4	918	3	US-10-369-800-1	Sequence 1, Appli
33	38	70.4	1056	3	US-09-252-991A-11806	Sequence 11806, A
34	38	70.4	1503	3	US-10-142-231-48	Sequence 48, Appl
35	38	70.4	1694	6	PCT-US95-04910-9	Sequence 9, Appli
36	38	70.4	3336	2	US-08-368-803-6	Sequence 6, Appli
C 37	38	70.4	3395	9	5223424-11	Patent No. 5223424
38	38	70.4	5153	6	PCT-US95-04910-8	Sequence 8, Appli
39	38	70.4	5426	2	US-08-663-566A-3	Sequence 3, Appli
40	38	70.4	5426	2	US-08-023-610-3	Sequence 3, Appli
41	38	70.4	5426	2	US-08-288-065A-3	Sequence 3, Appli
42	38	70.4	5426	2	US-08-362-240A-3	Sequence 1, Appli
43	38	70.4	5426	3	US-08-804-372A-1	Sequence 1, Appli
44	38	70.4	5426	6	PCT-US95-10245-3	Sequence 3, Appli
45	38	70.4	6220	3	US-09-600-087-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-949-016-154321  
; Sequence 154321, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 154321  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-154321

Alignment Scores: 91.4 Length: 601  
Pred. No.: 41.00 Matches: 7  
Score: 100.0% Conservative: 1  
Percent Similarity: 87.5% Mismatches: 0  
Best Local Similarity: 75.9% Indels: 0  
Query Match: 3 Gaps: 0  
DB:

US-09-581-651D-41 (1-10) x US-09-949-016-154321 (1-601)

QY 2 SerIleProProArgAsnLeuGly 9  
Db 117 AGCATACCACCCAGGAATATGGG 140

## RESULT 2

US-09-949-016-154322  
; Sequence 154322, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 154322  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-154322

Alignment Scores:  
Pred. No.: 91.4 Length: 601  
Score: 41.00 Matches: 7  
Percent Similarity: 100.0% Conservative: 1  
Best Local Similarity: 87.5% Mismatches: 0  
Query Match: 75.9% Indels: 0  
Gaps: 0

US-09-581-651D-41 (1-10) x US-09-949-016-154322 (1-601)

QY 2 SerIleProProArgAsnLeuGly 9  
Db 32 AGCATACCACCCAGGAATATGGG 55

## RESULT 3

US-09-949-016-154323  
; Sequence 154323, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 154323  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-154323

Alignment Scores:  
Pred. No.: 91.4 Length: 601  
Score: 41.00 Matches: 7

Percent Similarity: 100.0% Conservative: 1  
Best Local Similarity: 87.5% Mismatches: 0  
Query Match: 75.9% Indels: 0  
Gaps: 0

US-09-581-651D-41 (1-10) x US-09-949-016-154323 (1-601)

QY 2 SerIleProProArgAsnLeuGly 9  
Db 6 AGCATACCACCCAGGAATATGGG 29

## RESULT 4

US-09-949-016-16065  
; Sequence 16065, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16065  
; LENGTH: 212139  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-16065

Alignment Scores:  
Pred. No.: 6.69e+04 Length: 212139  
Score: 41.00 Matches: 7  
Percent Similarity: 100.0% Conservative: 1  
Best Local Similarity: 87.5% Mismatches: 0  
Query Match: 75.9% Indels: 0  
Gaps: 0

US-09-581-651D-41 (1-10) x US-09-949-016-16065 (1-212139)

QY 2 SerIleProProArgAsnLeuGly 9  
Db 150438 AGCATACCACCCAGGAATATGGG 150461

## RESULT 5

US-09-621-976-19038/c  
; Sequence 19038, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 19038  
; LENGTH: 426  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-621-976-19038

Alignment Scores:  
Pred. No.: 95.6 Length: 426  
Score: 40.00 Matches: 6

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 12, 2006, 10:25:20 ; Search time 144.264 Seconds  
(without alignments)  
4864.742 Million cell updates/sec

Title: US-09-581-651D-41

Perfect score: 54

Sequence: 1 VSIPPNLGY 10

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/abs/ABSSWEB.spool/US09581651/runat\_10052006\_181209\_176/app\_query.fasta.1  
-DB=EST -OPMT=fastap -SUFFIX=p2n.rst -MINMATCH=0 -LOOPEL=0 -LOPEXT=0  
-UNITS=bites -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs03h  
-USER=US09581651 @CGN 1.1 6731 @runat\_10052006\_181209\_176 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: \*  
1: gb\_est1: \*  
2: gb\_est2: \*  
3: gb\_est3: \*  
4: gb\_est4: \*  
5: gb\_est5: \*  
6: gb\_est6: \*  
7: gb\_est7: \*  
8: gb\_est8: \*  
9: gb\_est9: \*  
10: gb\_est10: \*  
11: gb\_est11: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	54	100.0	362	1	AI263888
C 2	54	100.0	391	1	AA234360
C 3	54	100.0	440	1	AI754331
C 4	54	100.0	670	5	BU620723
C 5	54	100.0	706	3	BM674187
C 6	54	100.0	747	6	CA423317
C 7	54	100.0	967	5	BX402381

C 8	54	100.0	1044	5	BX398837
C 9	54	100.0	1052	5	BX380582
C 10	46	85.2	238	2	BF757788
C 11	45	83.3	216	8	DR950135
C 12	45	83.3	333	8	DR926919
C 13	45	83.3	455	1	AA797190
C 14	45	83.3	716	8	DR951381
C 15	45	83.3	723	5	BX092161
C 16	45	83.3	778	8	DR923528
C 17	45	83.3	800	8	DR947777
C 18	45	83.3	913	9	DR647139
C 19	44	81.5	890	9	CC480099
C 20	44	81.5	935	10	AG882568
C 21	43	79.6	188	8	DN498580
C 22	43	79.6	237	7	CV191597
C 23	43	79.6	282	1	AV079208
C 24	43	79.6	363	3	BP634715
C 25	43	79.6	442	7	CN963693
C 26	43	79.6	482	9	B46912
C 27	43	79.6	591	9	BZ191710
C 28	43	79.6	653	9	AZ566286
C 29	43	79.6	684	10	AG102718
C 30	43	79.6	756	3	BQ217931
C 31	43	79.6	785	6	CB990018
C 32	43	79.6	811	8	DR838506
C 33	43	79.6	814	8	CX795219
C 34	43	79.6	865	10	AG834589
C 35	43	79.6	866	7	CJ029932
C 36	43	79.6	869	8	CX408686
C 37	43	79.6	873	8	DR451282
C 38	43	79.6	896	10	CL722789
C 39	43	79.6	905	8	DR467514
C 40	43	79.6	909	8	DN026258
C 41	43	79.6	923	7	CF993900
C 42	43	79.6	923	8	DR472075
C 43	43	79.6	945	5	BQ885747
C 44	43	79.6	962	10	CL137507
C 45	43	79.6	1169	8	DR142066

ALIGNMENTS

AI263888 362 bp mRNA linear EST 13-NOV-1998  
gi08d11.x1 Soares NHMPu S1 Homo sapiens cdna clone IMAGE:1855893  
3' similar to gb:X02761\_cds1 FIBRONECTIN PRECURSOR (HUMAN);, mRNA  
sequence.  
AI263888.1 GI:3872091  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 362)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Seq primer: -40UP from Gibco.  
Location/Qualifiers  
1. .362  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1855893"  
/tissue\_type="Pooled human melanocyte, fetal heart, and

```
pregnant uterus"
/lab_host="DH10B"
/clone_lib="Soares_NhMPu_S1"
/note="Organ: mixed (see below); Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBHM, pregnant uterus
NBHPU, and fetal heart NBHH19W) were mixed, and as circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."

ORIGIN
Alignment Scores:
Pred. No.: 9.66 Length: 362
Score: 54.00 Matches: 10
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 1 Gaps: 0

US-09-581-651D-41 (1-10) x AI263888 (1-362)
QY 1 ValSerIleProProArgAsnLeuGlyTyr 10
Db 184 GTGAGTATCCCAACCTTGGATAC 155

RESULT 2
AA234360/c
LOCUS AA234360
DEFINITION zr72c02.s1 Soares_NhMPu_S1 Homo sapiens cDNA clone IMAGE:668930 3'
similar to gb:X02761_cds1 FIBRONECTIN PRECURSOR (HUMAN);, mRNA
sequence.
ACCESSION AA234360.1 GI:1858635
VERSION AA234360
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 391)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-Merck EST Project 1997
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 678 Std Error: 0.00
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 376.
Location/Qualifiers
1..391
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:5562897"
/db_xref="taxon:9606"
/clone="IMAGE:668930"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/clone_lib="Soares_NhMPu_S1"

/note="Organ: mixed (see below); Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBHM, pregnant uterus
NBHPU, and fetal heart NBHH19W) were mixed, and as circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."

ORIGIN
Alignment Scores:
Pred. No.: 10.5 Length: 391
Score: 54.00 Matches: 10
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 1 Gaps: 0

US-09-581-651D-41 (1-10) x AA234360 (1-391)
QY 1 ValSerIleProProArgAsnLeuGlyTyr 10
Db 175 GTGAGTATCCCAACCTTGGATAC 146

RESULT 3
AI754331/c
LOCUS AI754331
DEFINITION cr23e08.x1 Human bone marrow stromal cells Homo sapiens cDNA clone
HBMSC cr23e08 3', mRNA sequence.
ACCESSION AI754331
VERSION AI754331.1 GI:5132595
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
REFERENCE 1 (bases 1 to 440)
AUTHORS Jia,L., Young,M.F., Powell,J., Yang,L., Ho,N.C., Hotchkiss,R.,
Robey,P.G. and Francomano,C.A.
Gene expression profile of human bone marrow stromal cells:
high-throughput expressed sequence tag sequencing analysis
Genomics 79 (1), 7-17 (2002)
11827452
Contact: Libin Jia
Medical Genetics Branch
National Human Genome Research Institute
10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
Tel: 301-402-4877
Fax: 301-496-7157
Email: libin@helix.nih.gov
DNA Sequencing and analyses by National Institutes of Health
Intramural Sequencing Center (NISC).
Plate: 23 row: e Column: 08
Seq primer: -21M13 forward primer (ABI).
Location/Qualifiers
1..440
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HBMSC_cr23e08"
/sex="mixed"
/tissue_type="bone marrow stroma"
/dev_stages="mixed"
/lab_host="XLI-Blue MRF"/SOLR"
/clone_lib="Human bone marrow stromal cells"
/note="Vector: pBluescript; Site1: EcoRI; Site 2: XhoI;
mRNA made from human bone marrow stroma, cDNA made by
oligo-dT priming. Directionally cloned. Size-selected for
average insert size >0.5 kb. Library constructed by Dr.
```



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 12, 2006, 09:52:22 ; Search time 16.1963 Seconds  
(without alignments)  
6172.414 Million cell updates/sec

Title: US-09-581-651d-41  
Perfect score: 54  
Sequence: 1 VSIPPRNLGY 10

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/abs/ABSSWEB.spool/US09581651/runat\_10052006\_181204\_141/app\_query.fasta.1  
-DB=N Geneseq -QFMT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs03p  
-USER=US09581651 @CGN\_1\_1\_1359 @runat\_10052006\_181204\_141 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_21.\*

- 1: Geneseqn1980s.\*
- 2: Geneseqn1990s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002as.\*
- 7: Geneseqn2002bs.\*
- 8: Geneseqn2003as.\*
- 9: Geneseqn2003bs.\*
- 10: Geneseqn2003cs.\*
- 11: Geneseqn2003ds.\*
- 12: Geneseqn2004as.\*
- 13: Geneseqn2004bs.\*
- 14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	100.0	1929	13	Adsl17488 Nucleotid
2	54	100.0	1929	13	Adr97657 Human fib
3	54	100.0	1929	14	Adw44479 Human fib
4	54	100.0	1929	14	Ady55703 Human fib

5	54	100.0	1929	14	ADY32457	Ady32457 Human fib
6	54	100.0	1929	14	AEBS6257	Aeb56257 Event seq
7	54	100.0	1929	14	AEBS78061	Aeb78061 Human DNA
8	54	100.0	2137	10	ADD18477	Add18477 Human pro
9	54	100.0	2147	2	AAx81299	Aax81299 Human mig
10	54	100.0	2358	13	ADR67201	Adr67201 Human bla
11	54	100.0	2402	14	ADZ26738	Adz26738 Human fib
12	54	100.0	2443	13	ADQ38575	Adq38575 Human SNP
13	54	100.0	2488	13	ADQ38581	Adq38581 Human SNP
14	54	100.0	4295	8	ACC72037	Acc72037 BCJ0770 g
15	54	100.0	11026	4	AAK72958	Aak72958 Human imm
16	54	100.0	31749	4	AAK72959	Aak72959 Human imm
17	54	100.0	78925	3	AAC89888	Aac89888 Human FN
18	45	83.3	79084	12	ADQ97563	Adq97563 Human can
19	44	81.5	28564	10	ADD47028	Add47028 Human gen
20	42	77.8	792	11	ACL32370	AcL32370 Rice abio
21	42	77.8	2234	2	AAZ00451	Aaz00451 Human sec
22	42	77.8	2234	10	ADA56384	Ada56384 Gene enco
23	42	77.8	8628	6	ABK64727	Abk64727 Human ben
24	42	77.8	8955	4	ABL24128	AbL24128 Drosophil
25	41	75.9	574	13	ADQ56307	Adq56307 Novel can
26	41	75.9	1110	4	ABL12307	AbL12307 Drosophil
27	41	75.9	2285	4	ABL20614	AbL20614 Drosophil
28	41	75.9	3330	4	ABL12306	AbL12306 Drosophil
29	41	75.9	228139	11	ACN44002	Acn44002 Human gen
30	40	74.1	480	10	ACD98232	AcD98232 Human col
31	40	74.1	662	6	ABQ57120	Abq57120 Human col
32	40	74.1	677	10	ADI02480	Adi02480 Human CDN
33	40	74.1	941	4	ABL09641	AbL09641 Drosophil
34	40	74.1	1480	10	ADB61974	AdB61974 Human CDN
35	40	74.1	1542	14	AEA18948	Aea18948 Methylo
36	40	74.1	1752	13	ADX63544	Adx63544 Plant ful
37	40	74.1	1906	8	ABX12015	Abx12015 Prostate
38	40	74.1	1930	6	ABS62728	Abs62728 Prostate
39	40	74.1	1930	12	ADL13228	AdL13228 Human ste
40	40	74.1	1930	14	ADV85752	Adv85752 Human CDN
41	40	74.1	1964	6	AAS94775	Aas94775 Human DNA
42	40	74.1	2288	4	ABL21423	AbL21423 Drosophil
43	40	74.1	2417	12	ADM91237	AdM91237 DNA homol
44	40	74.1	2417	14	ADM13961	Adw13961 Mouse CDN
45	40	74.1	2943	4	ABL18172	AbL18172 Drosophil

ALIGNMENTS

RESULT 1	
ADSL17488	
ID	ADSL17488 standard; DNA; 1929 BP.
XX	
AC	ADSL17488;
XX	
DT	02-DRC-2004 (first entry)
XX	
DE	Nucleotide sequence of human fibronectin 1.
XX	
KW	cell state; time-lapse profile; protein-protein interaction;
KW	drug screening; cancer; infectious disease; allergy; hypertension;
KW	hyperlipaemia; diabetes; cardiac disease; cerebral infarction; dementia;
KW	obesity; arteriosclerosis; infertility; mental disease; nervous disease;
KW	cataract; progeria; hypersensitivity; ultraviolet radiation; human;
KW	fibronectin 1; actin acting substance; transfection array; gene; ds.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..1929
FT	/*tag= a
FT	/product= "fibronectin 1"
XX	
PN	WO2004079007-A2.
XX	
PD	16-SEP-2004.
XX	

PF 03-MAR-2004; 2004WO-JP002694.  
XX  
XX PR  
XX 04-MAR-2003; 2003JP-00057870.  
XX  
XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.  
XX  
XX Miyake M, Yoshikawa T, Uchimura E, Miyake J;  
PI WPI; 2004-662438/64.  
DR P-PSDB; ADS17489.  
XX  
XX Presenting a state of a cell, useful for diagnosing and treating a  
PT disease, e.g. cancer, infectious disease, allergy, diabetes, dementia,  
PT obesity, infertility, or cataract, comprises obtaining a time-lapse  
PT profile of the cell.  
XX  
XX Disclosure; SEQ ID NO 1; 532pp; English.  
XX  
XX The specification describes a method and system for accurately presenting  
CC a state of a cell. The method comprises obtaining a time-lapse profile of  
CC the cell by time-lapse monitoring of a gene state associated with at  
CC least one gene derived from the cell, and presenting the time-lapse  
CC profile. The gene comprises a transcription control sequence, and the  
CC gene state includes expression of the gene. The method and system are  
CC useful presenting a state of a cell. The method can allow the elucidation  
CC of key protein-protein interactions suitable for targeting by drug  
CC screening protocols. The method is useful for diagnosing or treating a  
CC disease, e.g. cancer, infectious disease due to viruses or bacteria,  
CC allergy, hypertension, hyperlipaemia, diabetes, cardiac disease, cerebral  
CC infarction, dementia, obesity, arteriosclerosis, infertility, mental and  
CC nervous diseases, cataract, progeria, or hypersensitivity to ultraviolet  
CC radiation. The present sequence encodes human fibronectin 1. Bovine  
CC fibronectin was used as a candidate for an actin acting substance. The  
CC actin acting substance was used with transfection reagents and amplified  
CC plasmid DNA in assays using transfection arrays, in the course of the  
XX invention.  
XX  
XX SQ Sequence 1929 BP; 507 A; 462 C; 544 G; 416 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 6.69 Length: 1929  
Score: 54.00 Matches: 10  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 13 Gaps: 0  
US-09-581-651D-41 (1-10) x ADS17488 (1-1929)  
QY 1 ValSerIleProProArgGlnLeuGlyTyr 10  
Db 1897 GTGAGTATCCCAACCCAGAAACCTTGGATAC 1926  
RESULT 2  
ADR97657  
ID ADR97657 standard; DNA; 1929 BP.  
XX  
XX ADR97657;  
XX  
XX 02-DEC-2004 (first entry)  
DT Human fibronectin 1 DNA, an actin acting substance SeqID 1.  
DE human; gene; ds; transfection efficiency; actin acting substance;  
XX extracellular matrix; fibronectin 1; gene introduction reagent.  
KW  
XX Homo sapiens.  
OS  
XX Key Location/Qualifiers  
FH 1..1929  
FT /\*tag= a  
FT /product= "Fibronectin protein"  
XX

PN WO2004079332-A2.  
XX  
XX PD 16-SEP-2004.  
XX  
XX 03-MAR-2004; 2004WO-JP002696.  
XX  
XX 04-MAR-2003; 2003JP-00057869.  
XX  
XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.  
XX  
XX Miyake M, Yoshikawa T, Uchimura E, Miyake J;  
PI WPI; 2004-677173/66.  
DR P-PSDB; ADR97658.  
XX  
XX Composition comprising an actin acting substance or an actin acting  
PT substance and a target substance, useful for increasing the efficiency of  
PT introducing a target substance into a cell and in cell biology or genetic  
PT engineering.  
XX  
XX Claim 4; SEQ ID NO 1; 347pp; English.  
XX  
XX This invention relates to a novel composition and method for increasing  
CC the efficiency of introducing a target substance into a cell.  
CC Specifically, it refers to the introduction of DNA (e.g. the gene of  
CC interest to be transfected), polypeptides, sugars or complexes thereof  
CC into a cell, and comprises an actin acting substance. The present  
CC invention describes the actin acting substance as an extracellular matrix  
CC protein, a variant or fragment thereof selected from fibronectin, laminin  
CC or vitronectin. The composition further comprises a gene introduction  
CC reagent selected from cationic polymers, cationic lipids, and calcium  
CC phosphate, as well as a gold colloid particle that is contacted with the  
CC cell. As such, the composition, kit, device or method is useful for  
CC increasing the efficiency of introducing a target substance into a cell  
CC and thus is useful in the fields of cell biology, genetic engineering and  
CC molecular biology. This polynucleotide sequence is the human fibronectin  
CC DNA sequence of the invention.  
XX  
XX SQ Sequence 1929 BP; 507 A; 462 C; 544 G; 416 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 6.69 Length: 1929  
Score: 54.00 Matches: 10  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 100.0% Indels: 0  
DB: 13 Gaps: 0  
US-09-581-651D-41 (1-10) x ADR97657 (1-1929)  
QY 1 ValSerIleProProArgGlnLeuGlyTyr 10  
Db 1897 GTGAGTATCCCAACCCAGAAACCTTGGATAC 1926  
RESULT 3  
ADM44479  
ID ADM44479 standard; DNA; 1929 BP.  
XX  
XX ADM44479;  
XX  
XX 24-MAR-2005 (first entry)  
DT Human fibronectin 1 DNA.  
DE cell transduction; nerves; cell adhesion; fibronectin 1; ds; gene.  
XX  
XX Homo sapiens.  
OS  
XX Key Location/Qualifiers  
FH 1..1929  
FT /\*tag= a  
FT /product= "fibronectin 1"  
XX

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 12, 2006, 10:13:41 ; Search time 184.709 Seconds  
(without alignments)  
4616.200 Million cell updates/sec

Title: US-09-581-651D-41  
Perfect score: 54  
Sequence: 1 VSIPPNLGY 10

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Zgapop 6.0 , Zgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 117666282

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/abs/ABSSWEB pool/US09581651/runat\_10052006\_181206\_153/app.query.fasta.1  
-DB=GenEmbl -OFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pcc -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs04  
-USER=US09581651 @CGN 1.1 5142 @runat\_10052006\_181206\_153 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DRV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_in:\*  
3: gb\_env:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pr:\*  
9: gb\_to:\*  
10: gb\_sts:\*  
11: gb\_sy:\*  
12: gb\_un:\*  
13: gb\_vi:\*  
14: gb\_htg:\*  
15: gb\_pl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	100.0	471	6 CQ697541	CQ697541 Sequence
2	54	100.0	1929	6 CQ871810	CQ871810 Sequence
3	54	100.0	1929	6 CQ871828	CQ871828 Sequence

4	54	100.0	2147	6	BD137021	BD137021 Polypepti
5	54	100.0	2147	6	AX003229	AX003229 Sequence
6	54	100.0	2147	8	HSA276395	HSA276395 Homo sapi
7	54	100.0	2192	8	HSA535086	HSA535086 Homo sapi
8	54	100.0	2358	6	CQ875358	CQ875358 Sequence
9	54	100.0	4321	8	HSM806214	HSM806214 Homo sapi
10	54	100.0	5984	8	AJ849445	AJ849445 Homo sapi
c 11	54	100.0	143947	8	AC012462	AC012462 Homo sapi
c 12	49	90.7	110000	14	AC106675	AC106675 Rattus no
c 13	47	87.0	244068	14	AC131345	AC131345 Rattus no
c 14	47	87.0	247046	14	AC119441	AC119441 Rattus no
c 15	46	85.2	178271	9	AC091785	AC091785 Genomic s
c 16	45	83.3	6001	5	SQUCARPSYN	L31362 Squalus aca
c 17	45	83.3	9017	8	AL137121	AL137121 Human DNA
c 18	45	83.3	95771	8	AC093891	AC093891 Homo sapi
c 19	45	83.3	139452	9	AL954675	AL954675 Mouse DNA
c 20	45	83.3	150488	8	AL359392	AL359392 Human DNA
c 21	45	83.3	176204	14	AC140804	AC140804 Homo sapi
c 22	45	83.3	181755	14	AC015758	AC015758 Homo sapi
c 23	45	83.3	194093	9	AC139037	AC139037 Mus muscu
c 24	45	83.3	220102	9	AC158152	AC158152 Mus muscu
c 25	45	83.3	236521	14	AC129274	AC129274 Rattus no
c 26	44	81.5	105011	14	AC096724	AC096724 Homo sapi
c 27	44	81.5	110000	1	EX936398	Continuation (38 o
c 28	44	81.5	156278	14	AC015607	AC015607 Homo sapi
c 29	44	81.5	163339	14	AC149785	AC149785 Otolenur
c 30	44	81.5	169649	8	AL133328	AL133328 Human DNA
c 31	44	81.5	171940	14	AC148514	AC148514 Papio anu
c 32	44	81.5	173146	8	AC082649	AC082649 Homo sapi
c 33	44	81.5	180594	9	AC112270	AC112270 Mus muscu
c 34	44	81.5	207362	14	AC156716	AC156716 Bos tauru
c 35	44	81.5	213625	14	AC120950	AC120950 Rattus no
c 36	44	81.5	216381	9	AC120138	AC120138 Mus muscu
c 37	44	81.5	218524	9	AC151721	AC151721 Mus muscu
c 38	44	81.5	224009	14	AC164301	AC164301 Mus muscu
c 39	44	81.5	228405	14	AC123473	AC123473 Rattus no
c 40	44	81.5	241510	14	AC094856	AC094856 Rattus no
c 41	44	81.5	246248	14	AC130847	AC130847 Rattus no
c 42	44	81.5	267540	14	AC107153	AC107153 Rattus no
c 43	43	79.6	1100	13	HIM287013	AJ287013 Human imm
c 44	43	79.6	101299	14	AC022423	AC022423 Homo sapi
c 45	43	79.6	110000	1	CP000082	Continuation (12 o

#### ALIGNMENTS

RESULT 1	CQ697541	Sequence	42467	from Patent WO02070737.	471 bp	DNA	linear	PAT 03-FEB-2004
LOCUS	CQ697541	Sequence	42467	from Patent WO02070737.	471 bp	DNA	linear	PAT 03-FEB-2004
DEFINITION	CQ697541	Sequence	42467	from Patent WO02070737.	471 bp	DNA	linear	PAT 03-FEB-2004
ACCESSION	CQ697541	Sequence	42467	from Patent WO02070737.	471 bp	DNA	linear	PAT 03-FEB-2004
VERSION	CQ697541.1	GI:42248968						
KEYWORDS								
SOURCE		Homo sapiens (human)						
ORGANISM		Homo sapiens						
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
		Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;						
		Hominidae; Homo.						
REFERENCE	1	Liew,C.C., Marshall,W.E. and Zhang,H.						
AUTHORS		Compositions and methods relating to osteoarthritis						
TITLE		Patent: WO 02070737-A 42467 12-SEP-2002;						
JOURNAL		Chondrogene Inc. (CA)						
FEATURES		Location/Qualifiers						
source		1..471						
		/organism="Homo sapiens"						
		/mol_type="unassigned DNA"						
		/db_xref="taxon:9606"						

#### ORIGIN

Alignment Scores: 0.523 Length: 471  
Pred. No.: 54.00 Matches: 10  
Score:

Percent Similarity: 100.0%  
Best Local Similarity: 100.0%  
Query Match: 100.0%  
DB: 6  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-581-651D-41 (1-10) x CQ697541 (1-471)

QY 1 ValserilleProProAraLeuGlyTyr 10  
Db 393 GTGAGTATCCACCCAGAAACCTTGGATAC 422

## RESULT 2

CQ871810  
LOCUS CQ871810 1929 bp DNA linear PAT 27-SEP-2004  
DEFINITION Sequence 1 from Patent WO2004079332.  
ACCESSION CQ871810

VERSION CQ871810.1 GI:52745842

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominae; Homo.

## REFERENCE

1 Miyake, M.A., Yoshikawa, T.A., Uchimura, E.A. and Miyake, J.A.

AUTHORS Composition and method for increasing efficiency of introduction of

TITLE Target substance into cell

JOURNAL Patent: WO 2004079332-A 1 16-SEP-2004;

National Institute of Advanced Industrial Science and Technology

(JP)

## FEATURES

source

1. .1929

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

1. .1929

/notes="unnamed protein product; fibronectin 1"

/codon\_start=1

/protein\_id="CAH56867.1"

/translation="MLRGPGLLLAVQLGTAVPSTGASKSKRQAQOMVQPSFVA

VSQKPGCYDNGKHQYQINQWERTYLGALVCTCYGSGRGNCSKPEAEETCFDKYT

GNTRYGVDTYERPKDSMIWDCICAGRGRICTIANRCHGGQSYKIGDTRRRPHET

GGVMLFCVLGNGKGMTCPIAEKCFDHAAGTSYVVGTEKPKYQGMWVDCITLGE

SGRITCTSRNCRNDQRTSYRIGDTWSKDNRLQICITGNGRGEWCKERTHSV

QITSSGSGFTDRAAVYQPPHPQPPYGHCVTDGSGVVSVMQWLKTKQGNKQMLCT

CLNGVSCQETAVTYGNSNGEPCVLPFTYNGRTDSTTSYEQDQKYSFCTDHTVL

VTRGNSGALCHFPFLYNNHYTDCTSEGRDNMKNWCTTQNTYDADQKFCFPMMA

HEEICTTNEGVMYRIGDQKDHGMHMRCTCVNGRGEWTCIAYSQRLDQCIYDDI

TYNVNDTFRHKEEGHMLNCTCFGQGRGWKCDPVDQCDSTGTFTYQIGDSWEKVVH

GVRVCCYCGRGIGWHQCLQTYPSSSGPVEVFTETPSPNHPIQWNAPOPSHIS

KYLWRPVSIPRNLGY"

ORIGIN

Alignment Scores:

Pred. No.: 2.15 Length: 1929

Score: 54.00 Matches: 10

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0

DB: 6 Gaps: 0

US-09-581-651D-41 (1-10) x CQ871810 (1-1929)

QY 1 ValserilleProProAraLeuGlyTyr 10

Db 1897 GTGAGTATCCACCCAGAAACCTTGGATAC 1926

## RESULT 3

CQ871828  
LOCUS CQ871828 1929 bp DNA linear PAT 27-SEP-2004  
DEFINITION Sequence 1 from Patent WO2004079007.  
ACCESSION CQ871828

VERSION CQ871828.1 GI:52745858  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominae; Homo.

REFERENCE

1 Miyake, M.A., Yoshikawa, T.A., Uchimura, E.A. and Miyake, J.A.

AUTHORS Time-lapse cell analysis method

TITLE Patent: WO 2004079007-A 1 16-SEP-2004;

National Institute of Advanced Industrial Science and Technology

(JP)

FEATURES

source

1. .1929

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

1. .1929

/notes="unnamed protein product; fibronectin 1"

/codon\_start=1

/protein\_id="CAH56870.1"

/translation="MLRGPGLLLAVQLGTAVPSTGASKSKRQAQOMVQPSFVA

VSQKPGCYDNGKHQYQINQWERTYLGALVCTCYGSGRGNCSKPEAEETCFDKYT

GNTRYGVDTYERPKDSMIWDCICAGRGRICTIANRCHGGQSYKIGDTRRRPHET

GGVMLFCVLGNGKGMTCPIAEKCFDHAAGTSYVVGTEKPKYQGMWVDCITLGE

SGRITCTSRNCRNDQRTSYRIGDTWSKDNRLQICITGNGRGEWCKERTHSV

QITSSGSGFTDRAAVYQPPHPQPPYGHCVTDGSGVVSVMQWLKTKQGNKQMLCT

CLNGVSCQETAVTYGNSNGEPCVLPFTYNGRTDSTTSYEQDQKYSFCTDHTVL

VTRGNSGALCHFPFLYNNHYTDCTSEGRDNMKNWCTTQNTYDADQKFCFPMMA

HEEICTTNEGVMYRIGDQKDHGMHMRCTCVNGRGEWTCIAYSQRLDQCIYDDI

TYNVNDTFRHKEEGHMLNCTCFGQGRGWKCDPVDQCDSTGTFTYQIGDSWEKVVH

GVRVCCYCGRGIGWHQCLQTYPSSSGPVEVFTETPSPNHPIQWNAPOPSHIS

KYLWRPVSIPRNLGY"

ORIGIN

Alignment Scores:

Pred. No.: 2.15 Length: 1929

Score: 54.00 Matches: 10

Percent Similarity: 100.0% Conservative: 0

Best Local Similarity: 100.0% Mismatches: 0

Query Match: 100.0% Indels: 0

DB: 6 Gaps: 0

US-09-581-651D-41 (1-10) x CQ871828 (1-1929)

QY 1 ValserilleProProAraLeuGlyTyr 10

Db 1897 GTGAGTATCCACCCAGAAACCTTGGATAC 1926

RESULT 4

BD137021

LOCUS BD137021

DEFINITION Polypeptides, polynucleotides and uses thereof.

ACCESSION BD137021

VERSION BD137021.1 GI:3231966

KEYWORDS JP 2002508179-A/1.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominae; Homo.

REFERENCE

1 (bases 1 to 2147)

AUTHORS Schor, S.L. and Schor, A.M.

TITLE Polypeptides, polynucleotides and uses thereof

JOURNAL Patent: JP 2002508179-A 1 19-MAR-2002;

UNIVERSITY OF DUNDEE

COMMENT OS Homo sapiens (human)

PN JP 2002508179-A/1

PD 19-MAR-2002

PF 15-DEC-1998 JP 2000539133

PR 16-DEC-1997 GB 9726539.1

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2006, 19:42:45 ; Search time 0.598159 Seconds  
(without alignments)  
784.888 Million cell updates/sec

Title: US-09-581-651D-41

Perfect score: 54

Sequence: 1 VSIPPRNLGY 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 250354 seqs, 46948837 residues

Total number of hits satisfying chosen parameters: 250354

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:

- 1: /SID55/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.1\*
- 2: /SID55/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.1\*
- 3: /SID55/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.1\*
- 4: /SID55/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.1\*
- 5: /SID55/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.1\*
- 6: /SID55/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.1\*
- 7: /SID55/ptodata/2/pubpaa/US03\_NEW\_PUB.pep.1\*
- 8: /SID55/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.1\*
- 9: /SID55/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.1\*
- 10: /SID55/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.1\*
- 11: /SID55/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.1\*
- 12: /SID55/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.1\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	100.0	642	9	US-10-995-561-631
2	54	100.0	657	9	US-10-995-561-622
3	54	100.0	657	11	US-11-193-561-27
4	54	100.0	657	11	US-11-193-771-27
5	54	100.0	657	11	US-11-193-789-27
6	54	100.0	657	11	US-11-193-806-27
7	54	100.0	657	11	US-11-193-857-27
8	39	72.2	454	11	US-11-188-298-2724
9	37	68.5	3482	11	US-11-087-099-2068
10	37	68.5	3488	11	US-11-087-099-9005
11	36	66.7	3488	11	US-11-188-298-11717
12	36	66.7	3487	11	US-11-087-099-9068
13	36	66.7	3487	11	US-11-087-099-10423
14	35	64.8	132	11	US-11-188-298-18528
15	35	64.8	327	11	US-11-096-568A-11342
16	35	64.8	342	11	US-11-188-298-14861
17	35	64.8	374	11	US-11-096-568A-11341
18	35	64.8	400	11	US-11-096-568A-11340
19	35	64.8	436	11	US-11-188-298-3852
20	35	64.8	438	11	US-11-188-298-4885
21	35	64.8	432	11	US-11-188-298-5798

22	35	64.8	492	11	US-11-188-298-14169	Sequence 14169, A
23	35	64.8	600	11	US-11-096-568A-32116	Sequence 32116, A
24	35	64.8	666	11	US-11-096-568A-32115	Sequence 32115, A
25	35	64.8	670	11	US-11-096-568A-32114	Sequence 32114, A
26	35	64.8	820	11	US-11-098-686-11029	Sequence 11029, A
27	34	63.0	99	11	US-11-079-463-7832	Sequence 7832, Ap
28	34	63.0	113	9	US-10-475-075-816	Sequence 816, App
29	34	63.0	144	7	US-09-978-360A-613	Sequence 613, App
30	34	63.0	165	7	US-09-978-360A-497	Sequence 497, App
31	34	63.0	208	11	US-11-072-512-2284	Sequence 2284, Ap
32	34	63.0	269	9	US-10-495-597-14	Sequence 14, Appl
33	34	63.0	291	11	US-11-020-602-227	Sequence 227, App
34	34	63.0	327	9	US-10-506-454-337	Sequence 337, App
35	34	63.0	339	9	US-10-995-076-1	Sequence 1, Appli
36	34	63.0	339	9	US-10-995-076-2	Sequence 2, Appli
37	34	63.0	361	11	US-11-096-568A-34318	Sequence 34318, A
38	34	63.0	390	11	US-11-096-568A-34317	Sequence 34317, A
39	34	63.0	422	11	US-11-096-568A-33852	Sequence 33852, A
40	34	63.0	423	11	US-11-096-568A-34316	Sequence 34316, A
41	34	63.0	427	9	US-10-454-437-96	Sequence 96, Appl
42	34	63.0	440	8	US-10-196-749-34	Sequence 34, Appl
43	34	63.0	440	9	US-10-194-487-34	Sequence 34, Appl
44	34	63.0	440	9	US-10-195-883-34	Sequence 34, Appl
45	34	63.0	440	9	US-10-195-888-34	Sequence 34, Appl

## ALIGNMENTS

RESULT 1  
US-10-995-561-631  
; Sequence 631, Application US/10995561  
; Publication No. US20050272054A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
; TITLE OF INVENTION: DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001559  
; CURRENT APPLICATION NUMBER: US/10/995,561  
; CURRENT FILING DATE: 2004-11-24  
; NUMBER OF SEQ ID NOS: 85702  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 631  
; LENGTH: 642  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-995-561-631

Query Match 100.0%; Score 54; DB 9; Length 642;  
Best Local Similarity 100.0%; Pred. No. 0.043;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	VSIPPRNLGY 10
Db	633	VSIPPRNLGY 642

## RESULT 2

US-10-995-561-622  
; Sequence 622, Application US/10995561  
; Publication No. US20050272054A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
; TITLE OF INVENTION: DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001559  
; CURRENT APPLICATION NUMBER: US/10/995,561  
; CURRENT FILING DATE: 2004-11-24  
; NUMBER OF SEQ ID NOS: 85702  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 622

```
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-622

Query Match      100.0%; Score 54; DB 9; Length 657;
Best Local Similarity 100.0%; Pred. No. 0.044;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VSIPRNLGY 10
Db      648 VSIPRNLGY 657

RESULT 3
US-11-193-561-27
; Sequence 27, Application US/11193561
; Publication No. US20060024757A1
; GENERAL INFORMATION:
; APPLICANT: Hussa, Robert
; APPLICANT: Shorter, Simon
; TITLE OF INVENTION: Detection of Oncofetal Fibronectin for Selection of Concepti
; FILE REFERENCE: 17101-080001/831
; CURRENT APPLICATION NUMBER: US/11/193,561
; CURRENT FILING DATE: 2005-07-29
; PRIOR APPLICATION NUMBER: 60/592,823
; PRIOR FILING DATE: 2004-07-30
; PRIOR APPLICATION NUMBER: 60/592,803
; PRIOR FILING DATE: 2004-07-30
; PRIOR APPLICATION NUMBER: 60/592,825
; PRIOR FILING DATE: 2004-07-30
; PRIOR APPLICATION NUMBER: 60/592,804
; PRIOR FILING DATE: 2004-07-30
; PRIOR APPLICATION NUMBER: 60/592,824
; PRIOR FILING DATE: 2004-07-30
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo Sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank NM 054034
; DATABASE ENTRY DATE: 2005-06-10
US-11-193-561-27

Query Match      100.0%; Score 54; DB 11; Length 657;
Best Local Similarity 100.0%; Pred. No. 0.044;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VSIPRNLGY 10
Db      648 VSIPRNLGY 657

RESULT 4
US-11-193-771-27
; Sequence 27, Application US/11193771
; Publication No. US20060024722A1
; GENERAL INFORMATION:
; APPLICANT: Fisher-Colbrie, Mark
; APPLICANT: Hickok, Durlin
; APPLICANT: LaPointe, Jerome P.
; TITLE OF INVENTION: Samples for Detection of Oncofetal Fibronectin and uses thereof
; FILE REFERENCE: 17101-027001/828
; CURRENT APPLICATION NUMBER: US/11/193,771
; CURRENT FILING DATE: 2005-07-29
; PRIOR APPLICATION NUMBER: 60/592,823
; PRIOR FILING DATE: 2004-07-30
; PRIOR APPLICATION NUMBER: 60/592,803
; PRIOR FILING DATE: 2004-07-30
; PRIOR APPLICATION NUMBER: 60/592,825
; PRIOR FILING DATE: 2004-07-30
```

```
; PRIOR APPLICATION NUMBER: 60/592,804
; PRIOR FILING DATE: 2004-07-30
; PRIOR APPLICATION NUMBER: 60/592,824
; PRIOR FILING DATE: 2004-07-30
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo Sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank NM 054034
; DATABASE ENTRY DATE: 2005-06-10
US-11-193-771-27

Query Match      100.0%; Score 54; DB 11; Length 657;
Best Local Similarity 100.0%; Pred. No. 0.044;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VSIPRNLGY 10
Db      648 VSIPRNLGY 657

RESULT 5
US-11-193-789-27
; Sequence 27, Application US/11193789
; Publication No. US20060024723A1
; GENERAL INFORMATION:
; APPLICANT: Hussa, Robert
; APPLICANT: Fisher-Colbrie, Mark
; APPLICANT: LaPointe, Jerome
; APPLICANT: Senyeli, Andrew
; APPLICANT: Shorter, Simon
; TITLE OF INVENTION: Methods for detecting Oncofetal Fibronectin
; FILE REFERENCE: 17101-029001/830
; CURRENT APPLICATION NUMBER: US/11/193,789
; CURRENT FILING DATE: 2005-07-29
; PRIOR APPLICATION NUMBER: 60/592,823
; PRIOR FILING DATE: 2004-07-30
; PRIOR APPLICATION NUMBER: 60/592,803
; PRIOR FILING DATE: 2004-07-30
; PRIOR APPLICATION NUMBER: 60/592,825
; PRIOR FILING DATE: 2004-07-30
; PRIOR APPLICATION NUMBER: 60/592,804
; PRIOR FILING DATE: 2004-07-30
; PRIOR APPLICATION NUMBER: 60/592,824
; PRIOR FILING DATE: 2004-07-30
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo Sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: GenBank NM 054034
; DATABASE ENTRY DATE: 2005-06-10
US-11-193-789-27

Query Match      100.0%; Score 54; DB 11; Length 657;
Best Local Similarity 100.0%; Pred. No. 0.044;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VSIPRNLGY 10
Db      648 VSIPRNLGY 657

RESULT 6
US-11-193-806-27
; Sequence 27, Application US/11193806
; Publication No. US20060024724A1
; GENERAL INFORMATION:
```

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2006, 19:41:10 ; Search time 3.78834 Seconds  
(without alignments)  
1102.934 Million cell updates/sec

Title: US-09-581-651D-41

Perfect score: 54

Sequence: 1 VSIPPRNLGY 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	100.0	642	4	US-10-741-601-354
2	54	100.0	642	5	US-10-741-600-1066
3	54	100.0	657	4	US-10-741-601-359
4	54	100.0	657	5	US-10-741-600-1072
5	38	70.4	70	4	US-10-425-115-215232
6	38	70.4	71	4	US-10-424-599-186842
7	38	70.4	89	4	US-10-437-963-179849
8	38	70.4	98	4	US-10-425-115-316330
9	38	70.4	118	4	US-10-425-115-313669
10	38	70.4	145	6	US-11-097-143-40935
11	38	70.4	279	4	US-10-369-800-2
12	38	70.4	291	4	US-10-250-727-4
13	38	70.4	291	4	US-10-250-824-4
14	38	70.4	500	4	US-10-142-231-61
15	38	70.4	500	4	US-10-356-153-61
16	38	70.4	500	5	US-10-884-115-61
17	37	68.5	40	3	US-09-973-278-198
18	37	68.5	41	3	US-09-983-802-211
19	37	68.5	41	3	US-09-984-490-211
20	37	68.5	145	4	US-10-437-963-193731
21	37	68.5	152	4	US-10-425-115-301041
22	37	68.5	485	4	US-10-142-231-88
23	37	68.5	485	4	US-10-356-153-88
24	37	68.5	485	5	US-10-884-115-88
25	37	68.5	485	5	US-10-732-923-1081
26	36	66.7	45	4	US-10-424-599-167062
27	36	66.7	150	5	US-10-450-763-36134

28	36	66.7	154	3	US-09-925-297-783	Sequence 783, Appl
29	36	66.7	216	5	US-10-965-898-1	Sequence 1, Appl
30	36	66.7	512	4	US-10-142-231-87	Sequence 87, Appl
31	36	66.7	512	4	US-10-356-153-87	Sequence 87, Appl
32	36	66.7	512	5	US-10-884-115-87	Sequence 87, Appl
33	36	66.7	578	5	US-10-450-763-32551	Sequence 32551, A
34	36	66.7	738	5	US-10-450-763-38284	Sequence 38284, A
35	36	66.7	771	4	US-10-437-963-189959	Sequence 189959, A
36	36	66.7	822	4	US-10-437-963-104983	Sequence 104983, A
37	36	66.7	1572	6	US-11-097-143-153396	Sequence 15396, A
38	35	64.8	74	4	US-10-425-115-353331	Sequence 353331, A
39	35	64.8	107	4	US-10-425-115-283878	Sequence 283878, A
40	35	64.8	110	4	US-10-437-963-118856	Sequence 118856, A
41	35	64.8	123	4	US-10-425-115-329379	Sequence 329379, A
42	35	64.8	137	4	US-10-437-963-143622	Sequence 143622, A
43	35	64.8	139	4	US-10-437-963-186450	Sequence 186450, A
44	35	64.8	147	4	US-10-437-963-143028	Sequence 143028, A
45	35	64.8	147	5	US-10-450-763-34424	Sequence 34424, A

## ALIGNMENTS

## RESULT 1

US-10-741-601-354  
; Sequence 354, Application US/10741601  
; Publication No. US20040166519A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: STENOSIS, METHODS OF DETECTION AND USES THEREOF  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 26415  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 354  
; LENGTH: 642  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-741-601-354

Query Match 100.0%; Score 54; DB 4; Length 642;  
Best Local Similarity 100.0%; Pred. No. 0.79;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSIPPRNLGY 10  
Db 633 VSIPPRNLGY 642

## RESULT 2

US-10-741-600-1066  
; Sequence 1066, Application US/10741600  
; Publication No. US20050026169A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 73997  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1066  
; LENGTH: 642  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-741-600-1066

Query Match 100.0%; Score 54; DB 5; Length 642;  
Best Local Similarity 100.0%; Pred. No. 0.79;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 VSIPPRNLGY 10
      |||||
Db      633 VSIPPRNLGY 642

RESULT 3
US-10-741-601-359
; Sequence 359, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 359
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-359

Query Match      100.0%; Score 54; DB 4; Length 657;
Best Local Similarity 100.0%; Pred. No. 0.81;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VSIPPRNLGY 10
      |||||
Db      648 VSIPPRNLGY 657

RESULT 4
US-10-741-600-1072
; Sequence 1072, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1072
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1072

Query Match      100.0%; Score 54; DB 5; Length 657;
Best Local Similarity 100.0%; Pred. No. 0.81;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VSIPPRNLGY 10
      |||||
Db      648 VSIPPRNLGY 657

RESULT 5
US-10-425-115-215232
; Sequence 215232, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: Plants
; CURRENT APPLICATION NUMBER: US/10-425-115-215232
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 179849

; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 215232
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_127889C.1.pep
US-10-425-115-215232

Query Match      70.4%; Score 38; DB 4; Length 70;
Best Local Similarity 77.8%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 VSIPPRNLG 9
      | |||||
Db      56 VKYPPRNLG 64

RESULT 6
US-10-424-599-186842
; Sequence 186842, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 186842
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_139730C.1.pep
US-10-424-599-186842

Query Match      70.4%; Score 38; DB 4; Length 71;
Best Local Similarity 85.7%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      4 PPRNLGY 10
      |||||
Db      12 PPRNFGY 18

RESULT 7
US-10-437-963-179849
; Sequence 179849, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 179849
```



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2006, 19:25:45 ; Search time 1.02761 Seconds  
(without alignments)  
804.545 Million cell updates/sec

Title: US-09-581-651D-41

Perfect score: 54

Sequence: 1 VSIPPRNLGY 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/1/iaa/5 COMB.pap.\*
- 2: /cgn2\_6/ptodata/1/iaa/6 COMB.pap.\*
- 3: /cgn2\_6/ptodata/1/iaa/H COMB.pap.\*
- 4: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pap.\*
- 5: /cgn2\_6/ptodata/1/iaa/RE COMB.pap.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	70.4	148	2	US-09-270-767-42039
2	38	70.4	279	2	US-10-369-800-2
3	38	70.4	351	2	US-09-252-991A-28377
4	38	70.4	500	2	US-10-142-231-61
5	37	68.5	40	2	US-09-973-278-198
6	37	68.5	41	2	US-09-227-357-211
7	37	68.5	485	2	US-10-142-231-88
8	36	66.7	269	2	US-09-949-016-6927
9	36	66.7	497	2	US-09-270-767-62287
10	36	66.7	512	2	US-10-142-231-87
11	36	66.7	729	2	US-09-949-016-11190
12	36	66.7	782	2	US-09-270-767-46675
13	35	64.8	16	2	US-09-809-464-5
14	35	64.8	28	1	US-08-479-275D-17
15	35	64.8	28	1	US-08-479-275D-25
16	35	64.8	28	1	US-08-488-271B-17
17	35	64.8	28	1	US-08-488-271B-25
18	35	64.8	274	2	US-09-252-991A-32312
19	35	64.8	290	2	US-09-107-433-4222
20	35	64.8	315	2	US-09-583-110-4592
21	35	64.8	424	2	US-09-270-767-37179
22	35	64.8	424	2	US-09-270-767-37179
23	35	64.8	492	2	US-10-142-231-60
24	35	64.8	509	2	US-10-142-231-67
25	35	64.8	543	2	US-09-351-150A-15
26	35	64.8	834	2	US-09-198-452A-287
27	35	64.8	839	2	US-09-438-185A-276

Sequence 2, Appli  
Sequence 4302, Ap  
Sequence 21559, A  
Sequence 7461, Ap  
Sequence 5060, Ap  
Sequence 694, App  
Sequence 694, App  
Sequence 4206, Ap  
Sequence 7059, Ap  
Sequence 32, Appl  
Sequence 26, Appl  
Sequence 41196, A  
Sequence 56412, A  
Sequence 20776, A  
Sequence 22849, A  
Sequence 42649, A  
Sequence 19, Appl  
Sequence 1, Appl

28 35 64.8 841 2 US-09-546-990-2  
29 35 64.8 886 2 US-09-543-681A-4302  
30 35 64.8 953 2 US-09-252-991A-21559  
31 34 63.0 72 2 US-09-543-681A-7461  
32 34 63.0 77 2 US-09-513-999C-5060  
33 34 63.0 109 2 US-09-205-258-694  
34 34 63.0 109 2 US-10-004-860-694  
35 34 63.0 113 2 US-09-621-976-4206  
36 34 63.0 124 2 US-09-513-999C-7059  
37 34 63.0 130 2 US-09-434-840-32  
38 34 63.0 157 2 US-09-364-230-26  
39 34 63.0 164 2 US-09-270-767-41196  
40 34 63.0 164 2 US-09-270-767-56412  
41 34 63.0 205 2 US-09-252-991A-20776  
42 34 63.0 208 2 US-10-104-047-2284  
43 34 63.0 227 2 US-09-270-767-42649  
44 34 63.0 269 2 US-09-630-250B-19  
45 34 63.0 269 2 US-09-937-919-1

#### ALIGNMENTS

##### RESULT 1

US-09-270-767-42039  
; Sequence 42039, Application US/09270767  
; Patent No. 6703491  
; GENERAL INFORMATION:  
; APPLICANT: Homburger et al.  
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*  
; FILE REFERENCE: File Reference: 7326-094  
; CURRENT APPLICATION NUMBER: US/09/270,767  
; CURRENT FILING DATE: 1999-03-17  
; NUMBER OF SEQ ID NOS: 62517  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 42039  
; LENGTH: 148  
; TYPE: PRT  
; ORGANISM: *Drosophila melanogaster*  
US-09-270-767-42039

Query Match 70.4%; Score 38; DB 2; Length 148;  
Best Local Similarity 77.8%; Pred. No. 23;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 VSIPPRNLG 9

Db 84 VSIPPPNMG 92

##### RESULT 2

US-10-369-800-2  
; Sequence 2, Application US/10369800  
; Patent No. 6897033  
; GENERAL INFORMATION:  
; APPLICANT: The Procter & Gamble Company  
; APPLICANT: Dawson, Thomas  
; APPLICANT: Deangelis, Yvonne  
; APPLICANT: Johnstone, Kevin  
; APPLICANT: Kaczvinsky, Joseph  
; APPLICANT: Saunders, Charles  
; APPLICANT: Walter, Richard  
; TITLE OF INVENTION: No. 6897033el Fungal Lipase  
; FILE REFERENCE: No. 6897033el Fungal Lipase  
; CURRENT APPLICATION NUMBER: US/10/369,800  
; CURRENT FILING DATE: 2003-02-19  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 2  
; LENGTH: 279  
; TYPE: PRT  
; ORGANISM: *malassezia globosa*  
US-10-369-800-2

Query Match 70.4%; Score 38; DB 2; Length 279;  
Best Local Similarity 66.7%; Pred. No. 44;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 SIPRPNLGY 10  
:|||||  
Db 207 TVPPRALGY 215

## RESULT 3

US-09-252-991A-28377  
; Sequence 28377, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 28377  
; LENGTH: 351  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-28377

Query Match 70.4%; Score 38; DB 2; Length 351;  
Best Local Similarity 60.0%; Pred. No. 56;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VSIPRPNLGY 10  
:|:|:|:|  
Db 341 LSLPPAHLGY 350

## RESULT 4

US-10-142-231-61  
; Sequence 61, Application US/10142231  
; Patent No. 6787343  
; GENERAL INFORMATION:  
; APPLICANT: Croteau, Rodney et al.  
; TITLE OF INVENTION: CYTOCHROME P450 OXYGENASES AND THEIR USES  
; FILE REFERENCE: 62773  
; CURRENT APPLICATION NUMBER: US/10/142,231  
; CURRENT FILING DATE: 2002-05-08  
; PRIOR APPLICATION NUMBER: 60/165,250  
; PRIOR FILING DATE: 1999-11-12  
; NUMBER OF SEQ ID NOS: 95  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 61  
; LENGTH: 500  
; TYPE: PRT  
; ORGANISM: Taxus cuspidata  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(500)  
; OTHER INFORMATION: "Xaa" equals any peptide  
US-10-142-231-61

Query Match 70.4%; Score 38; DB 2; Length 500;  
Best Local Similarity 60.0%; Pred. No. 81;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VSIPRPNLGY 10  
|:|:|:|:|  
Db 52 VTLPFGNGLGF 61

RESULT 5  
US-09-973-278-198  
; Sequence 198, Application US/09973278  
; Patent No. 6924354  
; GENERAL INFORMATION:  
; APPLICANT: Fischer et al.  
; TITLE OF INVENTION: 123 Human Secreted Proteins  
; FILE REFERENCE: P2010P2  
; CURRENT APPLICATION NUMBER: US/09/973,278  
; CURRENT FILING DATE: 2001-10-10  
; PRIOR APPLICATION NUMBER: 60/239,899  
; PRIOR FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 09/227,357  
; PRIOR FILING DATE: 1999-01-08  
; PRIOR APPLICATION NUMBER: PCT/US98/13684  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/051,926  
; PRIOR FILING DATE: 1997-07-08  
; PRIOR APPLICATION NUMBER: 60/052,793  
; PRIOR FILING DATE: 1997-07-08  
; PRIOR APPLICATION NUMBER: 60/051,925  
; PRIOR FILING DATE: 1997-07-08  
; PRIOR APPLICATION NUMBER: 60/051,929  
; PRIOR FILING DATE: 1997-07-08  
; PRIOR APPLICATION NUMBER: 60/052,803  
; PRIOR FILING DATE: 1997-07-08  
; PRIOR APPLICATION NUMBER: 60/052,732  
; PRIOR FILING DATE: 1997-07-08  
; PRIOR APPLICATION NUMBER: 60/051,931  
; PRIOR FILING DATE: 1997-07-08  
; PRIOR APPLICATION NUMBER: 60/051,932  
; PRIOR FILING DATE: 1997-07-08  
; PRIOR APPLICATION NUMBER: 60/051,916  
; PRIOR FILING DATE: 1997-07-08  
; PRIOR APPLICATION NUMBER: 60/051,930  
; PRIOR FILING DATE: 1997-07-08  
; PRIOR APPLICATION NUMBER: 60/051,918  
; PRIOR FILING DATE: 1997-07-08  
; PRIOR APPLICATION NUMBER: 60/051,920  
; PRIOR FILING DATE: 1997-07-08  
; PRIOR APPLICATION NUMBER: 60/052,733  
; PRIOR FILING DATE: 1997-07-08  
; PRIOR APPLICATION NUMBER: 60/052,795  
; PRIOR FILING DATE: 1997-07-08  
; PRIOR APPLICATION NUMBER: 60/051,919  
; PRIOR FILING DATE: 1997-07-08  
; PRIOR APPLICATION NUMBER: 60/051,928  
; PRIOR FILING DATE: 1997-07-08  
; PRIOR APPLICATION NUMBER: 60/055,722  
; PRIOR FILING DATE: 1997-08-18  
; PRIOR APPLICATION NUMBER: 60/055,723  
; PRIOR FILING DATE: 1997-08-18  
; PRIOR APPLICATION NUMBER: 60/055,948  
; PRIOR FILING DATE: 1997-08-18  
; PRIOR APPLICATION NUMBER: 60/055,949  
; PRIOR FILING DATE: 1997-08-18  
; PRIOR APPLICATION NUMBER: 60/055,953  
; PRIOR FILING DATE: 1997-08-18  
; PRIOR APPLICATION NUMBER: 60/055,950  
; PRIOR FILING DATE: 1997-08-18  
; PRIOR APPLICATION NUMBER: 60/055,947  
; PRIOR FILING DATE: 1997-08-18  
; PRIOR APPLICATION NUMBER: 60/055,964  
; PRIOR FILING DATE: 1997-08-18  
; PRIOR APPLICATION NUMBER: 60/056,360  
; PRIOR FILING DATE: 1997-08-18  
; PRIOR APPLICATION NUMBER: 60/055,684  
; PRIOR FILING DATE: 1997-08-18  
; PRIOR APPLICATION NUMBER: 60/055,984  
; PRIOR FILING DATE: 1997-08-18  
; PRIOR APPLICATION NUMBER: 60/055,954  
; PRIOR FILING DATE: 1997-08-18

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2006, 19:13:45 ; Search time 4.72393 Seconds  
(without alignments)  
1493.521 Million cell updates/sec

Title: US-09-581-651d-41  
Perfect score: 54  
Sequence: 1 VSIPPRNLGY 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_05\_80.\*  
1: uniprot\_sprot.\*  
2: uniprot\_crembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	100.0	657	2	Q564H7 HUMAN
2	54	100.0	749	2	Q72391 HOMO SAPIEN
3	42	77.8	775	2	Q4HYV4 GIBBERELLA
4	41	75.9	181	2	Q56F00 AEROMONAS P
5	41	75.9	2515	2	Q77365 PLASMODIUM
6	40	74.1	267	2	Q77365 GIARDIA LAM
7	40	74.1	887	1	Q10213 SCHIZOSACCH
8	40	74.1	3689	2	Q7PPF9 ANOPHELES G
9	39	72.2	236	2	Q650E9 BACTEROIDES
10	39	72.2	331	2	Q65TX9 MANSUM
11	39	72.2	342	2	Q7SF17 NEUCR
12	39	72.2	411	2	Q4NJ95 ARTHROBACTE
13	39	72.2	454	2	Q8RP88 RHOO
14	39	72.2	644	2	Q4S6N4 TETNG
15	39	72.2	1708	2	Q7PVR6 DROSOPHILA
16	38	70.4	145	2	Q9VE17 DROME
17	38	70.4	189	2	Q8BBK5 SHEPC
18	38	70.4	283	2	Q8YBB3 BRUCELLA ME
19	38	70.4	303	2	Q869A2 ACHTE
20	38	70.4	306	2	Q576Y0 BRUCELLA AB
21	38	70.4	306	2	Q8FX28 BRUSU
22	38	70.4	315	2	Q8U7R8 AGRT5
23	38	70.4	337	1	TRUD MANSUM
24	38	70.4	364	2	Q5E0T4 VIBF1
25	38	70.4	824	2	Q7UA91 SYNEX
26	38	70.4	829	2	Q7V9E4 PROMM
27	38	70.4	849	2	Q4FLZ5 GRICK
28	38	70.4	877	2	Q98RX7 GUILLARDIA
29	38	70.4	920	1	PARC SYN3
30	38	70.4	1127	1	NUL133 CAEEL
31	37	68.5	186	2	Q84RE9 ARABIDOPSIS

32	37	68.5	243	2	Q6MKQ8 BDEBA	Q6mkq8 bdellovibri
33	37	68.5	297	2	Q9ZV59 ARATH	Q9zv59 arabidopsis
34	37	68.5	321	2	Q7FLJ2 ORYSA	Q7flj2 oryza sativ
35	37	68.5	324	2	Q84RF0 ARATH	Q84rf0 arabidopsis
36	37	68.5	325	2	Q84RF1 ARATH	Q84rf1 arabidopsis
37	37	68.5	372	2	Q9HGV6 ASCO	Q9hgv6 saitoella c
38	37	68.5	391	2	Q6G034 BARQU	Q6g034 bartonella
39	37	68.5	485	1	T13H TAXCU	Q8w4t9 taxus cuspi
40	37	68.5	485	2	Q6G2I3 TAXCU	Q6gz13 taxus cuspi
41	37	68.5	485	2	Q5BU48 CONCI	Q5bu48 taxus x med
42	37	68.5	485	2	Q56GD5 TAXCH	Q56gd5 taxus chine
43	37	68.5	515	2	Q5JDO0 PYRKO	Q5jdo0 pyrococcus
44	37	68.5	624	2	Q4SMS4 TETNG	Q4sms4 tetraodon n
45	37	68.5	792	2	Q6EPG4 ORYSA	Q6epg4 oryza sativ

#### ALIGNMENTS

RESULT 1  
Q564H7 HUMAN  
ID Q564H7 HUMAN PRELIMINARY; PRT; 657 AA.  
AC Q564H7;  
DT 10-MAY-2005 (TrEMBLrel. 30, Created)  
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)  
DE Migration stimulating factor.  
GN Name=FN1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;  
OC Homo.  
RX NCBI\_TaxID=9606;  
RN [1]  
RP NUCLEOTIDE SEQUENCE  
RA Kay R.A., Ellis I.R., Jones S.J., Perrier S., Florence M.M.,  
RA Schor A.W., Schor S.L.;  
RT "mRNA 3'-UTR truncation: a novel mechanism regulating gene  
expression";  
RL Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.  
DR SMR; Q564H7.48-140, 183-275, 305-464.  
DR Ensembl; ENSG00000115414; Homo sapiens.  
DR GO; GO:0005576; C:extracellular region; IEA.  
DR InterPro; IPR006209; EGF like.  
DR InterPro; IPR000083; Fibrinctn1.  
DR InterPro; IPR000562; FN\_type2\_col\_bd.  
DR InterPro; IPR001724; Glyco\_hydro\_58.  
DR Pfam; PF00039; fn1; 9.  
DR Pfam; PF00040; fn2; 2.  
DR PRINTS; PR00012; FNTYPEI.  
DR PRINTS; PR00013; FNTYPEII.  
DR PRINTS; PR00849; GLHYDLASE58.  
DR SMART; SM00058; FN1; 9.  
DR SMART; SM00059; FN2; 2.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE; PS01253; FN1\_1; 9.  
DR PROSITE; PS1091; FN1\_2; 9.  
DR PROSITE; PS00023; FN2\_1; 2.  
DR PROSITE; PS1092; FN2\_2; 2.  
KW Repeat.

SQ SEQUENCE 657 AA; 73752 MW; 8CA04E64486ABCD0 CRC64;  
Query Match 100.0%; Score 54; DB 2; Length 657;  
Best Local Similarity 100.0%; Pred. No. 0.23;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VSIPPRNLGY 10  
Db 648 VSIPPRNLGY 657

RESULT 2

```

Q72391 HUMAN PRELIMINARY; PRT; 749 AA.
AC Q72391;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZ568B18150.
GN Name=DKFZp686B18150;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Human colon endothelial primary cell culture;
RA Bloecher H., Boecher M., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538045; CAD97984.1; -; mRNA.
DR HSSP; Q96KP7; 1PBR.
DR SMR; Q72391; 140-232, 275-367, 397-556.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0008243; F:plasma membrane activator activity; IEA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR000083; Fibrinctn1.
DR InterPro; IPR000562; FN Type II.
DR InterPro; IPR001724; Glyco_hydro_58.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00849; GLYDRLASE58.
DR ProDom; PD000995; FN_Type_II_2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR PROSITE; PS00022; EGF_1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein
SQ SEQUENCE 749 AA; 83524 MW; C8DDF97F3ED2F0DE CRC64;

Query Match 100.0%; Score 54; DB 2; Length 749;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSIPPRNLGY 10
DB 740 VSIPPRNLGY 749

RESULT 3
Q4HYV4 GIBZE
ID Q4HYV4_GIBZE PRELIMINARY; PRT; 775 AA.
AC Q4HYV4;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=FG09854.1;
OS Gibberella zeae PH-1.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
OX NCBI_TaxID=229533;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PH-1;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Arachchi H.M., Barna N., Bastien V., Bloom T., Boguslavskiy L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collamore A., Cook A., Cooke P., Corum B., Dearellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., FitzGerald M., Gage D., Galagan J.,

```

```

RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kanal M., Kamat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., Maclean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mlenga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neil D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramaamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schuback R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander B.;
RT Fusarium graminearum genome sequence."
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAC01000407; EAA76021.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 775 AA; 85797 MW; 02EFC46C826B4002 CRC64;

Query Match 77.8%; Score 42; DB 2; Length 775;
Best Local Similarity 77.8%; Pred. No. 54;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 SIPPRNLGY 10
DB 397 SYPPKNLGY 405

RESULT 4
Q56F00 9CAUD
ID Q56F00_9CAUD PRELIMINARY; PRT; 181 AA.
AC Q56F00;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein PHG31ORF011c.
GN Name=PHG31ORF011c; ORFNames=PHG31p11;
OS Aeromonas phage 31.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX NCBI_TaxID=321023;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Nolan J.M., Petrov V., Bertrand C., Krusch H.M., Karam J.D.;
RT "Comparative analysis of the Aeromonas bacteriophage 31 genome."
RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY962392; AAX63500.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 181 AA; 20612 MW; 21783CD9C2F8C497 CRC64;

Query Match 75.9%; Score 41; DB 2; Length 181;
Best Local Similarity 60.0%; Pred. No. 17;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VSIPPRNLGY 10
DB 89 VVVPFRNWGF 98

RESULT 5
Q77365 PLAF7
ID Q77365_PLAF7 PRELIMINARY; PRT; 2515 AA.
AC Q77365;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein MAL3P4.16.
GN Name=MAL3P4.16; Synonyms=PRC0485w;

```

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2006, 19:20:36 ; Search time 0.766871 Seconds  
(without alignments)  
1254.667 Million cell updates/sec

Title: US-09-581-651D-41

Perfect score: 54

Sequence: 1 VSIPPRNLGY 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR1.\*

2: PIR2.\*

3: PIR3.\*

4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41	75.9	2523	2 T18477	hypothetical prote
2	40	74.1	887	2 T38885	probable ATP-depen
3	38	70.4	283	2 AB3633	nirv precursor [im
4	38	70.4	315	2 A99192	nirv precursor (AF
5	38	70.4	315	2 AI3094	nitrite reductase,
6	38	70.4	877	2 D90086	DNA gyrase A-subun
7	38	70.4	944	2 S75188	DNA topoisomerase
8	38	70.4	1102	2 S44772	C29E4.4 protein -
9	37	68.5	297	2 E84731	hypothetical prote
10	37	68.5	1345	2 S55669	tegument protein 7
11	36	66.7	312	2 T23492	hypothetical prote
12	36	66.7	340	2 T23492	hypothetical prote
13	36	66.7	494	2 T32644	hypothetical prote
14	36	66.7	746	2 AG2039	serine/threonine k
15	36	66.7	1136	2 T30878	dynein heavy chain
16	35.5	65.7	484	2 G86475	unknown protein, 1
17	35	64.8	163	2 H72759	probable dCTP deam
18	35	64.8	238	2 E87055	probable conserved
19	35	64.8	251	2 S72791	hypothetical prote
20	35	64.8	271	2 B83617	hypothetical prote
21	35	64.8	310	2 T19379	hypothetical prote
22	35	64.8	468	2 S70992	DNA topoisomerase
23	35	64.8	662	2 I36792	hypothetical prote
24	35	64.8	669	2 S46519	luciferin-binding
25	35	64.8	822	2 B98009	DNA topoisomerase
26	35	64.8	822	2 D95141	DNA gyrase chain A
27	35	64.8	822	2 T43720	DNA topoisomerase
28	35	64.8	829	2 G86763	DNA topoisomerase
29	35	64.8	829	2 T46556	DNA topoisomerase

30	35	64.8	830	2 G96900	DNA gyrase (topois
31	35	64.8	833	2 H81700	DNA gyrase, chain
32	35	64.8	834	2 B86525	DNA gyrase subunit
33	35	64.8	834	2 G72098	DNA gyrase chain A
34	35	64.8	834	2 F81571	DNA gyrase, chain
35	35	64.8	835	2 AH1861	DNA topoisomerase
36	35	64.8	836	2 F71546	probable DNA gyras
37	35	64.8	858	2 S30571	DNA topoisomerase
38	35	64.8	860	2 S76534	hypothetical prote
39	35	64.8	872	2 AB1914	DNA gyrase A chain
40	35	64.8	878	2 T43767	DNA topoisomerase
41	35	64.8	883	2 T03577	DNA topoisomerase
42	35	64.8	894	2 E82221	DNA gyrase, chain
43	35	64.8	902	2 AF3362	DNA topoisomerase
44	35	64.8	905	2 A97734	DNA topoisomerase
45	35	64.8	905	2 H71731	DNA gyrase chain A

ALIGNMENTS

RESULT 1

T18477

hypothetical protein C0485w - malaria parasite (Plasmodium falciparum)

C;Species: Plasmodium falciparum

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C;Accession: T18477

R;Lawson, D.; Bowman, S.; Barrell, B.

submitted to the EMBL Data Library, November 1998

A;Reference number: Z18937

A;Accession: T18477

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-2523 <LAW>

A;Cross-references: UNIPROT:O77365; UNIPARC:UPI000017CC31; EMBL:AL008970; NID:el407852;

C;Genetics:

A;Map position: 3

A;Introns: 148/3

A;Note: C0485w

Query Match 75.9%; Score 41; DB 2; Length 2523;  
Best Local Similarity 60.0%; Pred. No. 45;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VSIPPRNLGY 10  
|.:|||:|  
Db 2263 VHVPPHNIGY 2272

RESULT 2

T38885

probable ATP-dependent DNA helicase (EC 3.6.1.1-) - fission yeast (Schizosaccharomyces po

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C;Accession: T38885

R;Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, February 1996

A;Reference number: Z21807

A;Accession: T38885

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-887 <MUR>

A;Cross-references: UNIPROT:Q10213; UNIPARC:UPI000013A195; EMBL:Z69380; PIDN:CAA93344.1;

A;Experimental source: strain 972h-; cosmid c4H3

C;Genetics:

A;Gene: SPDB:SPAC4H3.05

A;Map position: 1

A;Introns: 121/2; 209/3; 353/2

C;Keywords: hydrolase

Query Match 74.1%; Score 40; DB 2; Length 887;  
Best Local Similarity 55.8%; Pred. No. 23;  
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 VSIPPRNLG 9
Db      431 INVPPRNIG 439

RESULT 3
AB3633
Nirv precursor [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AB3633
R:DelVecchio, V.G.; Kaprat, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A:Reference number: AD3252; PMID:11756688
A:Accession: AB3633
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-283 <KUR>
A:CROSS-references: UNIPROT:Q8YBB3; UNIPARC:UPI00000586C8; GB:AE008918; PIDN:AAL54229.1;
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME110987
A:Map position: 11

Query Match      70.4%; Score 38; DB 2; Length 283;
Best Local Similarity 60.0%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 VSIPPRNLG 10
Db      253 VGVPDPNLGF 262

RESULT 4
A99192
Nirv precursor (AF040987) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C:Species: Agrobacterium tumefaciens
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: A99192
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; PMID:21608551; PMID:11743194
A:Accession: A99192
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-315 <KUR>
A:CROSS-references: UNIPROT:Q8U7R8; UNIPARC:UPI00000D254E; GB:AE007870; PIDN:AAK89059.1;
C:Genetics:
A:Gene: AGR_L_972
A:Map position: linear chromosome

Query Match      70.4%; Score 38; DB 2; Length 315;
Best Local Similarity 60.0%; Pred. No. 17;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 VSIPPRNLG 10
Db      283 VGVPDPNLGF 292

RESULT 5
AI3094
Nitrite reductase, Nirv precursor [imported] - Agrobacterium tumefaciens (strain C58, Du
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AI3094
R:Wood, D.W.; Stetzel, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell

```

```

; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AI3094
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-315 <KUR>
A:CROSS-references: UNIPROT:Q8U7R8; UNIPARC:UPI00000D254E; GB:AE008689; PIDN:AAL45175.1;
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: nirv
A:Map position: linear chromosome

Query Match      70.4%; Score 38; DB 2; Length 315;
Best Local Similarity 60.0%; Pred. No. 17;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 VSIPPRNLG 10
Db      283 VGVPDPNLGF 292

RESULT 6
D90086
DNA gyrase A-subunit [imported] - Guillardia theta nucleomorph
C:Species: nucleomorph Guillardia theta
A:Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 05-Oct-2004
C:Accession: D90086
R:Douglas, S.; Zauner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Reit
Nature 410, 1091-1096, 2001
A:Title: The highly reduced genome of an enslaved algal nucleus.
A:Reference number: A99082; MUID:11323671; PMID:11323671
A:Accession: D90086
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-877 <DOU>
A:CROSS-references: UNIPROT:Q98RX7; UNIPARC:UPI000008C0E5; GB:AF165818; NID:g13794448; P
C:Genetics:
A:Gene: gyra
A:Map position: 1
A:Genome: nucleomorph
C:Superfamily: Type II topoisomerase, subunit A; phage T4 DNA topoisomerase (ATP-hydroly
C:Keywords: nucleomorph

Query Match      70.4%; Score 38; DB 2; Length 877;
Best Local Similarity 87.5%; Pred. No. 53;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 SIPPRNLG 9
Db      239 SIPPHNLG 246

RESULT 7
S75188
DNA topoisomerase (ATP-hydrolyzing) (EC 5.99.1.3) chain A - Synechocystis sp. (strain PC
N:Alternate names: DNA gyrase; protein sll1941
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 05-Oct-2004
C:Accession: S75188
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S75188
A>Status: nucleic acid sequence not shown; translation not shown

```

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2006, 19:13:10 ; Search time 381.064 Seconds  
(without alignments)  
740.245 Million cell updates/sec

Title: US-09-581-651d-2

Perfect score: 3687

Sequence: 1 MLRGPFGLLLLLAVQCLGTA.....ISKYILRWRPVSPRNLGY 642

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_21.\*

1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*  
9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3687	100.0	660	2	AAY28901 Human mig
2	3680	99.8	642	8	Adr67316 Human bla
3	3680	99.8	642	8	Adsl7489 Amino aci
4	3680	99.8	642	8	Adr97658 Human fib
5	3680	99.8	642	9	Adw44480 Human fib
6	3680	99.8	642	9	Ady55704 Human fib
7	3680	99.8	642	9	Ady32458 Human fib
8	3680	99.8	642	9	AEB56258 Event seq
9	3680	99.8	642	9	Aeb78062 Human fib
10	3664	99.4	642	8	Adq39403 Human myo
11	3653.5	99.1	657	9	Adz26739 Human fib
12	3646.5	98.9	657	8	Adq39409 Human myo
13	3599.5	97.6	2176	9	Adz26741 Human fib
14	3599.5	97.6	2182	8	Adr66462 Human pro
15	3599.5	97.6	2182	8	Adr66120 Human pro
16	3599.5	97.6	2220	6	ABO01289 Human cel
17	3599.5	97.6	2286	6	ABR40124 Human cel
18	3599.5	97.6	2236	9	Adz26745 Human fib
19	3599.5	97.6	2330	9	Adz26749 Human fib
20	3599.5	97.6	2355	6	ABR58335 NM_00202
21	3599.5	97.6	2355	7	Adp65196 Human fib
22	3599.5	97.6	2355	8	Adg89560 Human fib
23	3599.5	97.6	2355	8	ADO55175 Protein #
24	3599.5	97.6	2355	8	Adq26085 Fibronect

25	3599.5	97.6	2355	8	ADQ29668 Human col
26	3599.5	97.6	2355	8	ADR67315 Human bla
27	3599.5	97.6	2355	9	ADV70223 Tumor-ase
28	3599.5	97.6	2355	9	ADZ26565 Human fib
29	3599.5	97.6	2355	9	AEA04490 Human pro
30	3599.5	97.6	2355	9	AEB17606 Human fib
31	3599.5	97.6	2355	9	AEB47084 Human fib
32	3599.5	97.6	2386	8	ADO55174 Protein #
33	3599.5	97.6	2386	8	ADU24084 Human thr
34	3599.5	97.6	2421	9	ADZ26743 Human fib
35	3595.5	97.5	2386	5	AAO17353 Human fib
36	3595.5	97.5	2386	6	ABR81866 Human fib
37	3595.5	97.5	2386	7	ADD18770 Human dis
38	3595.5	97.5	2386	7	ADE63324 Human pro
39	3595.5	97.5	2386	8	ADR99200 Fibronect
40	3595.5	97.5	2386	9	ADW63879 Human fib
41	3595.5	97.5	2386	9	ADX57710 Rheumatoi
42	3595.5	97.5	2386	9	ADX69339 Human hep
43	3595.5	97.5	2386	9	ADZ36436 Human fib
44	3594.5	97.5	2476	9	ADZ26747 Human fib
45	3593.5	97.5	2446	3	AAB50377 Human fib

#### ALIGNMENTS

##### RESULT 1

AAY28901

ID AAY28901 standard; protein; 660 AA.

XX AAY28901;

DT 21-SEP-1999 (first entry)

DE Human migration stimulating factor (MSF) 1-alpha protein.

XX Migration stimulatory factor; MSF; cell migration; modulation; human;

KW wound healing; scarring; MSF1-alpha.

XX Homo sapiens.

XX WO9931233-A1.

XX 24-JUN-1999.

XX 15-DEC-1998; 98WO-GB003766.

XX 16-DEC-1997; 97GB-00026539.

XX (UYDU-) UNIV DUNDEE.

XX Schor SL, Schor AM;

XX WPI; 1999-430039/36.

XX N-PSDB; AAX81299.

PT Proteins with cell migration stimulatory activity used in treating wound and preventing scarring.

XX Claim 1; Page 53; 86pp; English.

XX The invention provides a human migration stimulatory factor (MSF) protein. Host cells containing a replicable vector comprising the MSF encoding nucleic acid can be used for the recombinant production of the protein. The polypeptide can be used for modulating cell migration, healing a wound and for preventing scarring. The present sequence represents the human MSF1-alpha protein

XX Sequence 660 AA;

Query Match 100.0%; Score 3687; DB 2; Length 660;

Best Local Similarity 100.0%; Pred. No. 1.4e-235;

Matches 642; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MLRGPGLLLAVQCLGTAVPSTGASKSRQAQMVQPSVAVSQSPGCVGDKGHYQ 60
DB 19 MLRGPGLLLAVQCLGTAVPSTGASKSRQAQMVQPSVAVSQSPGCVGDKGHYQ 78
QY 61 INQWERTYLGALVCTCYGSGRGFNCESKPEAEETCFDKYTGNTYRVGDTYERPKDSMI 120
DB 79 INQWERTYLGALVCTCYGSGRGFNCESKPEAEETCFDKYTGNTYRVGDTYERPKDSMI 138
QY 121 WDCCTCAGGRISCTIANRCHGGOSYKIGDTRWRPHETGGYMLCEVCLGNGKGWTK 180
DB 139 WDCCTCAGGRISCTIANRCHGGOSYKIGDTRWRPHETGGYMLCEVCLGNGKGWTK 198
QY 181 PIAEKCFDHAAGTSYVVGTEWEPYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 240
DB 199 PIAEKCFDHAAGTSYVVGTEWEPYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 258
QY 241 RIGDTSKKNRGNLLQICITGNRGKWCERHTSVQTTSSGSPFTDVRAAVYQPPHP 300
DB 259 RIGDTSKKNRGNLLQICITGNRGKWCERHTSVQTTSSGSPFTDVRAAVYQPPHP 318
QY 301 QPPYGHCVTDGSGVYVSGMWLKTQGNKQMLCTCLGNGVSCQETAVTQTYGNSGEP 360
DB 319 QPPYGHCVTDGSGVYVSGMWLKTQGNKQMLCTCLGNGVSCQETAVTQTYGNSGEP 378
QY 361 VLPFTYNDRTDSTTSNYEQDKYSFCTDHTVLVQTRGNSNGALCHFPFLYNNHYTDCT 420
DB 379 VLPFTYNDRTDSTTSNYEQDKYSFCTDHTVLVQTRGNSNGALCHFPFLYNNHYTDCT 438
QY 421 SEGRDNMKWCGTQNTYDADQKFGFCPMAAHEICTTNEGVMYRIGDQWDKQHDGMHMR 480
DB 439 SEGRDNMKWCGTQNTYDADQKFGFCPMAAHEICTTNEGVMYRIGDQWDKQHDGMHMR 498
QY 481 CTCVNGRGWTCIAYSQLRDQCIYDDITVYNDTFHKBHEGHMLNCTCFQGRGRWKC 540
DB 499 CTCVNGRGWTCIAYSQLRDQCIYDDITVYNDTFHKBHEGHMLNCTCFQGRGRWKC 558
QY 541 DPVDCQDSETGTFTYQIGDSWEKYVHGVRQCYCYGRGIGEWHCQPLQTYPSSSGPVEVF 600
DB 559 DPVDCQDSETGTFTYQIGDSWEKYVHGVRQCYCYGRGIGEWHCQPLQTYPSSSGPVEVF 618
QY 601 ITETSPQNSHPIQWNAPOPSHISKYLRLWRPVSIPPRNLGY 642
DB 619 ITETSPQNSHPIQWNAPOPSHISKYLRLWRPVSIPPRNLGY 660
```

## RESULT 2

ADR67316  
ID ADR67316 standard; protein; 642 AA.

AC ADR67316;

DT 02-DEC-2004 (first entry)

DE Human bladder cancer associated amino acid sequence.

KW bladder cancer tissue; bladder cancer; cytostatic.

OS Homo sapiens.

PN WO2004076613-A2.

XX 10-SEP-2004.

XX 24-FEB-2004; 2004WO-DE000364.

PR 26-FEB-2003; 2003DE-01009729.

XX (HERR/) HERR A.

PA (HINZ/) HINZMANN B.

PA (DAHL/) DAHL E.

PA (STAUB/) STAUB E.

PA (PILAR/) PILARSKY C.

(SPEC/) SPECHT T.

XX Herr A, Hinzmann B, Dahl E, Staub E, Pilarsky C, Specht T;

XX WPI; 2004-653385/63.

XX New nucleic acids, and encoded proteins, from bladder cancer tissue,  
PT useful for diagnosis, treatment and in screening for specific binding  
PT agents.

XX Claim 2; Fig 3; 112pp; German.

XX The present invention describes nucleic acids (I) associated with bladder  
CC cancer tissue. Also described: (1) peptides and proteins (II) containing  
CC an amino acid sequence encoded by (I); (2) a method for diagnosing  
CC bladder cancer (BC), or monitoring its progression, that uses (I), (II)  
CC or agents (Z) that inhibit, or bind to, (I) or (II), labelled with a  
CC reporter; and (3) a method for treating BC that uses (I), (II) or (Z).  
CC (I) and (II) have cytostatic activity. (I) and (II) can be used to detect  
CC (and monitor progression of) bladder cancer (BC), or the risk of  
CC developing it; to screen for specific binding agents (Z), and to treat  
CC BC. (Z) are also useful as diagnostic and therapeutic agents. The present  
CC sequence represents a human amino acid sequence associated with bladder  
CC cancer, which is used in the exemplification of the present invention.

XX Sequence 642 AA;

Query Match 99.8%; Score 3680; DB 8; Length 642;

Best Local Similarity 99.8%; Pred. No. 3.8e-235;

Matches 641; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLRGPGLLLAVQCLGTAVPSTGASKSRQAQMVQPSVAVSQSPGCVGDKGHYQ 60

DB 1 MLRGPGLLLAVQCLGTAVPSTGASKSRQAQMVQPSVAVSQSPGCVGDKGHYQ 60

QY 61 INQWERTYLGALVCTCYGSGRGFNCESKPEAEETCFDKYTGNTYRVGDTYERPKDSMI 120

DB 61 INQWERTYLGALVCTCYGSGRGFNCESKPEAEETCFDKYTGNTYRVGDTYERPKDSMI 120

QY 121 WDCCTCAGGRISCTIANRCHGGOSYKIGDTRWRPHETGGYMLCEVCLGNGKGWTK 180

DB 121 WDCCTCAGGRISCTIANRCHGGOSYKIGDTRWRPHETGGYMLCEVCLGNGKGWTK 180

QY 181 PIAEKCFDHAAGTSYVVGTEWEPYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 240

DB 181 PIAEKCFDHAAGTSYVVGTEWEPYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 240

QY 241 RIGDTSKKNRGNLLQICITGNRGKWCERHTSVQTTSSGSPFTDVRAAVYQPPHP 300

DB 241 RIGDTSKKNRGNLLQICITGNRGKWCERHTSVQTTSSGSPFTDVRAAVYQPPHP 300

QY 301 QPPYGHCVTDGSGVYVSGMWLKTQGNKQMLCTCLGNGVSCQETAVTQTYGNSGEP 360

DB 301 QPPYGHCVTDGSGVYVSGMWLKTQGNKQMLCTCLGNGVSCQETAVTQTYGNSGEP 360

QY 361 VLPFTYNDRTDSTTSNYEQDKYSFCTDHTVLVQTRGNSNGALCHFPFLYNNHYTDCT 420

DB 361 VLPFTYNDRTDSTTSNYEQDKYSFCTDHTVLVQTRGNSNGALCHFPFLYNNHYTDCT 420

QY 421 SEGRDNMKWCGTQNTYDADQKFGFCPMAAHEICTTNEGVMYRIGDQWDKQHDGMHMR 480

DB 421 SEGRDNMKWCGTQNTYDADQKFGFCPMAAHEICTTNEGVMYRIGDQWDKQHDGMHMR 480

QY 481 CTCVNGRGWTCIAYSQLRDQCIYDDITVYNDTFHKBHEGHMLNCTCFQGRGRWKC 540

DB 481 CTCVNGRGWTCIAYSQLRDQCIYDDITVYNDTFHKBHEGHMLNCTCFQGRGRWKC 540

QY 541 DPVDCQDSETGTFTYQIGDSWEKYVHGVRQCYCYGRGIGEWHCQPLQTYPSSSGPVEVF 600

DB 541 DPVDCQDSETGTFTYQIGDSWEKYVHGVRQCYCYGRGIGEWHCQPLQTYPSSSGPVEVF 600

QY 601 ITETSPQNSHPIQWNAPOPSHISKYLRLWRPVSIPPRNLGY 642

|||||



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 12, 2006, 11:45:28 ; Search time 4776.6 Seconds  
(without alignments)  
822.262 Million cell updates/sec

Title: US-09-581-651D-2  
Perfect score: 3687  
Sequence: 1 MLRGPGLLLAVQCLGTA.....ISKYLWRPVSTPRNLGY 642

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 9312410 seqs, 2039259788 residues  
Total number of hits satisfying chosen parameters: 18624820

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/abs/ABSSWEB.spool/US09581651/runat\_10052006\_181225\_634/app\_query.fasta.1  
-DB=Published Applications NA New -QFMT=fastap -SUFFIX=p2n.rnpbn -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -HOST=abss06p  
-USER=US09581651 @CGN 1.1490 @runat\_10052006\_181225\_634 -NCPUs=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMSOULT=120  
-WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New:  
1: /SIDSS/ptodata/2/pubpna/US08 NEW PUB.seq1.\*  
2: /SIDSS/ptodata/2/pubpna/US06 NEW PUB.seq.\*  
3: /SIDSS/ptodata/2/pubpna/US07 NEW PUB.seq.\*  
4: /SIDSS/ptodata/2/pubpna/US08 NEW PUB.seq.\*  
5: /SIDSS/ptodata/2/pubpna/US09 NEW PUB.seq.\*  
6: /SIDSS/ptodata/2/pubpna/US09 NEW PUB.seq.\*  
7: /SIDSS/ptodata/2/pubpna/US09 NEW PUB.seq.\*  
8: /SIDSS/ptodata/2/pubpna/US09 NEW PUB.seq.\*  
9: /SIDSS/ptodata/2/pubpna/US10 NEW PUB.seq.\*  
10: /SIDSS/ptodata/2/pubpna/US10 NEW PUB.seq.\*  
11: /SIDSS/ptodata/2/pubpna/US10 NEW PUB.seq.\*  
12: /SIDSS/ptodata/2/pubpna/US10 NEW PUB.seq.\*  
13: /SIDSS/ptodata/2/pubpna/US10 NEW PUB.seq.\*  
14: /SIDSS/ptodata/2/pubpna/US11 NEW PUB.seq.\*  
15: /SIDSS/ptodata/2/pubpna/US11 NEW PUB.seq.\*  
16: /SIDSS/ptodata/2/pubpna/US11 NEW PUB.seq.\*  
17: /SIDSS/ptodata/2/pubpna/US11 NEW PUB.seq.\*  
18: /SIDSS/ptodata/2/pubpna/US11 NEW PUB.seq.\*  
19: /SIDSS/ptodata/2/pubpna/US60 NEW PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result Query

Alignment Scores: 5.07e-281 Length: 2443  
Pred. No.: 2443

No.	Score	Match	Length	DB	ID	Description
1	3658	99.2	2443	10	US-10-995-561-114	Sequence 114, App
2	3653.5	99.1	2402	17	US-11-193-561-26	Sequence 26, Appl
3	3653.5	99.1	2402	17	US-11-193-771-26	Sequence 26, Appl
4	3653.5	99.1	2402	17	US-11-193-789-26	Sequence 26, Appl
5	3653.5	99.1	2402	17	US-11-193-806-26	Sequence 26, Appl
6	3653.5	99.1	2402	17	US-11-193-857-26	Sequence 26, Appl
7	3640.5	98.7	2488	10	US-10-995-561-105	Sequence 105, App
8	3599.5	97.6	7912	17	US-11-193-561-24	Sequence 24, Appl
9	3599.5	97.6	7912	17	US-11-193-771-24	Sequence 24, Appl
10	3599.5	97.6	7912	17	US-11-193-789-24	Sequence 24, Appl
11	3599.5	97.6	7912	17	US-11-193-806-24	Sequence 24, Appl
12	3599.5	97.6	7912	17	US-11-193-857-24	Sequence 24, Appl
13	3599.5	97.6	8232	10	US-10-821-234-693	Sequence 693, App
14	3599.5	97.6	8272	17	US-11-193-561-22	Sequence 22, Appl
15	3599.5	97.6	8272	17	US-11-193-771-22	Sequence 22, Appl
16	3599.5	97.6	8272	17	US-11-193-789-22	Sequence 22, Appl
17	3599.5	97.6	8272	17	US-11-193-806-22	Sequence 22, Appl
18	3599.5	97.6	8272	17	US-11-193-857-22	Sequence 22, Appl
19	3599.5	97.6	8374	17	US-11-193-561-20	Sequence 20, Appl
20	3599.5	97.6	8374	17	US-11-193-771-20	Sequence 20, Appl
21	3599.5	97.6	8374	17	US-11-193-789-20	Sequence 20, Appl
22	3599.5	97.6	8374	17	US-11-193-806-20	Sequence 20, Appl
23	3599.5	97.6	8374	17	US-11-193-857-20	Sequence 20, Appl
24	3599.5	97.6	8449	17	US-11-193-561-18	Sequence 18, Appl
25	3599.5	97.6	8449	17	US-11-193-771-18	Sequence 18, Appl
26	3599.5	97.6	8449	17	US-11-193-789-18	Sequence 18, Appl
27	3599.5	97.6	8449	17	US-11-193-806-18	Sequence 18, Appl
28	3599.5	97.6	8449	17	US-11-193-857-18	Sequence 18, Appl
29	3599.5	97.6	8647	17	US-11-193-561-16	Sequence 16, Appl
30	3599.5	97.6	8647	17	US-11-193-771-16	Sequence 16, Appl
31	3599.5	97.6	8647	17	US-11-193-789-16	Sequence 16, Appl
32	3599.5	97.6	8647	17	US-11-193-806-16	Sequence 16, Appl
33	3599.5	97.6	8647	17	US-11-193-857-16	Sequence 16, Appl
34	3599.5	97.6	8815	17	US-11-193-561-14	Sequence 14, Appl
35	3599.5	97.6	8815	17	US-11-193-771-14	Sequence 14, Appl
36	3599.5	97.6	8815	17	US-11-193-789-14	Sequence 14, Appl
37	3599.5	97.6	8815	17	US-11-193-806-14	Sequence 14, Appl
38	3599.5	97.6	8815	17	US-11-193-857-14	Sequence 14, Appl
39	3586.5	97.3	6510	10	US-10-995-561-112	Sequence 112, App
40	3586.5	97.3	7823	10	US-10-995-561-117	Sequence 117, App
41	3586.5	97.3	7935	10	US-10-995-561-111	Sequence 111, App
42	3586.5	97.3	7935	10	US-10-995-561-113	Sequence 113, App
43	3586.5	97.3	7959	10	US-10-995-561-108	Sequence 108, App
44	3586.5	97.3	8013	10	US-10-995-561-104	Sequence 104, App
45	3586.5	97.3	8155	10	US-10-995-561-116	Sequence 116, App

## ALIGNMENTS

RESULT 1  
US-10-995-561-114  
; Sequence 114, Application US/10995561  
; Publication No. US20050272054A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
; TITLE OF INVENTION: DETECTION AND USES THEREOF  
; FILE REFERENCE: C0001559  
; CURRENT APPLICATION NUMBER: US/10/995,561  
; CURRENT FILING DATE: 2004-11-24  
; NUMBER OF SEQ ID NOS: 85702  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 114  
; LENGTH: 2443  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-995-561-114

Score:	3658.00	Matches:	638
Percent Similarity:	99.4%	Conservative:	0
Best Local Similarity:	99.4%	Mismatches:	4
Query Match:	99.2%	Indels:	0
DB:	10	Gaps:	0
US-09-581-651D-2 (1-642) x US-10-995-561-114 (1-2443)			
QY	1	MetLeuArgGlyProGlyProGlyLeuLeuLeuAlaValGlnCysLeuGlyThrAla	20
DB	373	ATGCTTAGGGGTCCGGGGCCGGGCTGCTGCTGGCCGTCWGTGCTGGGGACAGCG	432
QY	21	ValProSerThrGlyAlaSerLysSerLysArgGlnAlaGlnMetValGlnProGln	40
DB	433	GTCCCTCCACGGGAGCTCGAAGCAAGAGCGAGGCTCAGCAATGGTTGAGCCCCAG	492
QY	41	SerProValAlaValSerGlnSerLysProGlyCysTyrAspAsnGlyLysHisTyrGln	60
DB	493	TCCCGGTGGCTGTCAGTCAAGCAAGCCGGTGTATTGACAAATGGAAAACTATCAG	552
QY	61	IleAsnGlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGly	80
DB	553	ATAAATCAACAGTGGGAGCGGACCTACCTAGGCAATGCGTTGGTTGTACTTGTATGA	612
QY	81	GlySerArgGlyPheAsnCysGluSerLysProGluAlaGluGluThrCysPheAspLys	100
DB	613	GGAAGCCGAGTTTAACTCGGAGAGTAACCTGAAGCTGAAGAGACTTGTCTTGACAAG	672
QY	101	TyrThrGlyAsnThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIle	120
DB	673	TACACTGGGAACAACCTTACCAGTGGGTGACACTTATGAGCGTCTTAAAGACTCCATGATC	732
QY	121	TrpAspCysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArg	140
DB	733	TGGGACTGTACCTGCAATYGGGCTGGCGAGGAGAAATAGCTGTACCATCGCAACCCG	792
QY	141	CysHisGluGlyGlyGlnSerTyrLysIleGlyAspThrTyrArgArgProHisGluThr	160
DB	793	TGCCATGAGGGGGTCAGTCTCTACAGATTGGTGACACCTGGAGGAGACCAATGAGACT	852
QY	161	GlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluTrpThrCysLys	180
DB	853	GGTGGTTACATGTTAGAGTGTGTGTCTTGGTAATGGAAAAAGAGATGGACCTGCAAG	912
QY	181	ProIleAlaGlnLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr	200
DB	913	CCCATAGCTGAGAAGTGTGTGATCATGTCTGGGACTTCTTATGTGTGGAGAAACG	972
QY	201	TrpGluLysProTyrGlnGlyTyrMetMetValAspCysThrCysLeuGlyGlySer	220
DB	973	TGGGAGAGCCCTACCAAGCTGGATGATGGTAGATTGTACTTGGCTGGGAGAGGACGC	1032
QY	221	GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr	240
DB	1033	GGACGCATCATTGCACTTCTAGAAATAGATGCAAGCATCAGGACCAAGGACATCCTAT	1092
QY	241	ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys	260
DB	1093	AGAAATGGAGACACCTGGAGCAAGAGGATAATCGAGGAAACCTGTCTCAGTGCATCTGC	1152
QY	261	ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer	280
DB	1153	ACAGGCNAAGCCGAGGAGGTGGAGTGTGAGAGGCACACCTCTGTGGAGACCACATCG	1212
QY	281	SerGlySerGlyProPheThrAspValArgAlaValTyrGlnProGlnProHisPro	300
DB	1213	AGCGGATCTGGCCCTTCCCGATGTTCTGTCAGACTGTTTACCAACCGCAGCCCTCACCC	1272
QY	301	GlnProProProTyrGlyHisCysValThrAspSerGlyValValTyrSerValGlyMet	320
DB	1273	CAGCCTCTCCCTATGGCCACTGTGTACAGACAGAGTGGTGTGTCTACTCTGTGGGGATG	1332
QY	321	GlnTrpLeuLysThrGlnGlyAsnLysGlnMetLeuCysThrCysLeuGlyAsnGlyVal	340
DB	1333	CAGTGGCTGAAGACACACAGGAATAAGCAATGCTTTGACGCTGCTGGCAACGAGTGC	1392
QY	341	SerCysGlnGluThrAlaValThrGlnThrTyrGlyGlyAsnSerAsnGlyGluProCys	360
DB	1393	AGCTGCCAAGACACAGCTGTAAACCCAGACTTACRGTTGGCACTCAAAATGGAGAGCATGT	1452
QY	361	ValLeuProPheThrTyrAsnAspArgThrAspSerThrThrSerAsnTyrGluGlnAsp	380
DB	1453	GTCTTACCATTACCTTACCAATGGCAGGAGCTGCAGCACAACCTTCGAATTTATGACAGGAC	1512
QY	381	GlnLysTyrSerPheCysThrAspHisThrValLeuValGlnThrArgGlyGlyAsnSer	400
DB	1513	CAGAAATACTCTTTCTGCACACACACTGTTTGGTTTCAGACTCGAGGAGGAATTC	1572
QY	401	AsnGlyAlaLeuCysHisPheProPheLeuTyrAsnAsnHisAsnTyrThrAspCysThr	420
DB	1573	AATGGTGCCTTGTGCACCTTCCCTTCTATACAAACACACAATTTACACTGATTGCACT	1632
QY	421	SerGluGlyArgArgAspAsnMetLysTyrCysGlyThrThrGlnAsnTyrAspAlaAsp	440
DB	1633	TCTGAGGGCAGAGAGACAAACATGAAGTGGTGTGGAGCCACACAGAACTATGATGCCGAC	1692
QY	441	GlnLysPheGlyPheCysProMetAlaAlaHisGluGluIleCysThrThrAsnGluGly	460
DB	1693	CAGAAAGTTTGGGTTCTGCCCATGGCTGCCACGAGGAATCTGCACACCAATGAAGG	1752
QY	461	ValMetTyrArgIleGlyAspGlnTrpAspLysGlnHisAspMetGlyHisMetMetArg	480
DB	1753	GTCACTGATCCGATTTGGAGATCAGTGGGATAAGCAGCATGATGGGTGCATGATGAGG	1812
QY	481	CysThrCysValGlyAsnGlyArgGlyGluTrpThrCysIleAlaTyrSerGlnLeuArg	500
DB	1813	TGCAGTGTGTTGGGAATGGTCTGGGGAATGGACATGCTTTCCTTCTCGCAGCTTCGA	1872
QY	501	AspGlnCysIleValAspAspIleThrTyrAsnValAsnAspThrPheHisLysArgHis	520
DB	1873	GATCAGTGCAATTTGATGACATCATTTACAATGTGAACGACACACATTCACAAAGCGTCAT	1932
QY	521	GluGluGlyHisMetLeuAsnCysThrCysPheGlyGlnGlyArgGlyArgTrpLysCys	540
DB	1933	GAAGAGGGGCACATGCTGAACCTGTATCATGCTTCGGTCAAGGGTGGGGCAGGTGGAAGTGT	1992
QY	541	AspProValAspGlnCysGlnAspSerGluThrGlyThrPheTyrGlnIleGlyAspSer	560
DB	1993	GATCCCGTCGACCAATGCCAGGATTCAGAGACTGGGACGTTTTATCAAAATGGAGATTCA	2052
QY	561	TrpGluLysTyrValHisGlyValArgTyrGlnCysTyrCysTyrArgGlyIleGly	580
DB	2053	TGGGAGAGATGTTGCTGCTGTCAGATACCACTGCTACTGCTATGGCCGTGGCATTGG	2112
QY	581	GluTrpHisCysGlnProLeuGlnThrTyrProSerSerSerGlyProValGluValPhe	600
DB	2113	GAGTGGCATGTCACACCTTTACAGACCTATCCAAAGCTCAAGTGGTCTCTGCGAAGTATTT	2172
QY	601	IleThrGluThrProSerGlnProAsnSerHisProIleGlnTrpAsnAlaProGlnPro	620
DB	2173	ATCACTGAGACTCCGAGTCAGCCCACTCCCAACCCCATCCAGTGGAAATGCACACAGCCA	2232
QY	621	SerHisIleSerLysTyrIleLeuArgTrpArgProValSerIleProProArgAsnLeu	640
DB	2233	TCTCACATTTCCAAAGTACATTTCTAGGTGGAGACCTGTGAGTATCCCAACCAACCTT	2292
QY	641	GlyTyr 642	
DB	2293	GGATAC 2298	

## RESULT 2

US-11-193-561-26

; Sequence 26, Application US/11193561

; Publication NO. US20060024757A1

; GENERAL INFORMATION:

; APPLICANT: Hussa, Robert

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 12, 2006, 11:24:42 ; Search time 1526.23 Seconds  
(without alignments)  
5217.711 Million cell updates/sec

Title: US-09-581-651D-2

Perfect score: 3687

Sequence: 1 MLKGPFGGLLLAVQCLGTA.....ISKYILRWPRVSIIPRNLGY 642

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O=/abs/ABSSWEB spool/US09581651/runat.10052006.181222.525/app.query.fasta.1  
-DB=Published Applications\_NA\_Main -QWTF=fastap -SUFFIX=p2n.rnpbm  
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1  
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext  
-HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs03p  
-USER=US09581651 @CGN 1.1 2064 @runat.10052006.181222.525 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -BSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA\_Main:

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3671	99.6	2127	9	US-10-956-157-4288
2	3665	99.4	2443	7	US-10-741-601-70
3	3665	99.4	2443	8	US-10-741-600-238
4	3664	99.4	2127	6	US-10-210-120-49
5	3664	99.4	2127	9	US-10-909-035-49
6	3647.5	98.9	2488	7	US-10-741-601-75
7	3647.5	98.9	2488	8	US-10-741-600-244

8	3599.5	97.6	7361	7	US-10-236-392-3
9	3599.5	97.6	8027	7	US-10-447-161-8
10	3599.5	97.6	8027	7	US-10-734-564-27
11	3599.5	97.6	8027	9	US-10-852-335A-53
12	3599.5	97.6	8027	9	US-10-287-436A-81
13	3599.5	97.6	8815	8	US-10-868-577A-62
14	3599.5	97.6	8815	8	US-10-868-549-21
15	3599.5	97.5	6510	7	US-10-741-601-72
16	3599.5	97.5	6510	8	US-10-741-600-241
17	3593.5	97.5	7823	7	US-10-741-601-77
18	3593.5	97.5	7823	8	US-10-741-600-245
19	3593.5	97.5	7848	7	US-10-741-601-78
20	3593.5	97.5	7848	8	US-10-741-600-246
21	3593.5	97.5	7935	7	US-10-741-601-74
22	3593.5	97.5	7935	7	US-10-741-600-240
23	3593.5	97.5	7959	7	US-10-741-601-81
24	3593.5	97.5	7959	8	US-10-741-600-249
25	3593.5	97.5	8013	7	US-10-741-601-71
26	3593.5	97.5	8013	8	US-10-741-600-242
27	3593.5	97.5	8155	7	US-10-741-601-79
28	3593.5	97.5	8155	8	US-10-741-600-247
29	3593.5	97.5	8226	7	US-10-741-601-69
30	3593.5	97.5	8226	8	US-10-741-600-237
31	3593.5	97.5	8278	7	US-10-741-601-82
32	3593.5	97.5	8278	8	US-10-741-600-250
33	3593.5	97.5	8332	7	US-10-741-601-73
34	3593.5	97.5	8332	8	US-10-741-600-239
35	3593.5	97.5	8371	7	US-10-741-601-76
36	3593.5	97.5	8371	8	US-10-741-600-243
37	3592.5	97.4	7795	5	US-10-084-817-2
38	3592.5	97.4	7867	5	US-10-098-841-6
39	3592.5	97.4	8044	6	US-10-240-965-121
40	3592.5	97.4	8044	9	US-10-765-700-135
41	3592.5	97.4	8062	5	US-10-098-841-5
42	3592.5	97.4	8137	5	US-10-098-841-8
43	3592.5	97.4	8230	5	US-10-098-841-7
44	3524	95.6	4295	6	US-10-144-194A-51
45	3524	95.6	4295	8	US-10-491-566-51

ALIGNMENTS

RESULT 1

US-10-956-157-4288  
; Sequence 4288, Application US/10956157  
; Publication No. US20050118625A1  
; GENERAL INFORMATION:  
; APPLICANT: Wyeth  
; APPLICANT: Mounts, William  
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH  
; TITLE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES  
; FILE REFERENCE: 031896-043000 (AM 101081)  
; CURRENT APPLICATION NUMBER: US/10/956,157  
; CURRENT FILING DATE: 2004-10-04  
; NUMBER OF SEQ ID NOS: 319805  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4288  
; LENGTH: 2127  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-10-956-157-4288

Alignment Scores:

Pred. No.:	0	Length:	2127
Score:	3671.00	Matches:	640
Percent Similarity:	99.7%	Conservative:	0
Best Local Similarity:	99.7%	Mismatches:	2
Query Match:	99.6%	Indels:	0
DB:	9	Gaps:	0

US-09-581-651D-2 (1-642) x US-10-956-157-4288 (1-2127)

Qy 1 MetLeuArgGlyProGlyLeuLeuLeuAlaValGlnCysLeuGlyThrAla 20



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 12, 2006, 10:33:26 ; Search time 412.574 Seconds  
(without alignments)  
4149.058 Million cell updates/sec

Title: US-09-581-651D-2

Perfect score: 3687

Sequence: 1 MLRGPGLLLAVQCLGTA.....ISKYILWRVPVSIPPNLGY 642

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/abs/ABSSWEB.spool/US09581651/runat.10052006.181213.253/app.query.fasta\_1  
-DB=Issued Patents NA -OFMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosun62 -TRANS-human40.cdi  
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-HOST=abs04 -USER=US09581651 @CGN 1.1.237 @runat.10052006.181213.253 -NCPU=6  
-ICPU=3 -NO MAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.\*

1: /cgn2\_6/ptodata/1/ina/1 COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5 COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/H COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/PCTUS COMB.seq.\*  
7: /cgn2\_6/ptodata/1/ina/PP COMB.seq.\*  
8: /cgn2\_6/ptodata/1/ina/RE COMB.seq.\*  
9: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	3592.5	97.4	8044	3	US-09-566-921-135
2	3578.5	97.1	7803	2	Sequence 135, App
3	3578.5	97.1	7803	6	Sequence 1, Appli
4	3448.5	93.5	7679	3	PCT-US93-12687-1
5	3448.5	93.5	7680	3	US-09-220-132-38
6	3448.5	93.5	7680	6	PCT-US93-655-1289
7	3433.5	93.1	7705	2	US-08-259-569-16
8	3433.5	93.1	7705	2	US-08-826-885-16
9	3428.5	93.0	7705	9	Patent No. 5455158-2

10	1665.5	45.2	986	2	US-07-637-250A-8	Sequence 8, Appli
11	1665.5	45.2	986	2	US-08-145-061-8	Sequence 8, Appli
12	325.5	8.8	2109	3	US-09-799-451-345	Sequence 345, App
13	311	8.4	2335	3	US-09-799-451-346	Sequence 346, App
14	306.5	8.3	2335	2	US-08-392-678-33	Sequence 33, Appl
15	305.5	8.3	2334	2	US-08-457-304A-33	Sequence 33, Appl
16	305.5	8.3	2334	2	US-08-456-701A-33	Sequence 33, Appl
17	305.5	8.3	2334	3	US-08-684-932A-33	Sequence 33, Appl
18	305.5	8.3	2334	3	US-09-023-655-996	Sequence 996, App
19	305.5	8.3	2334	3	US-09-949-016-704	Sequence 704, App
20	305.5	8.3	2335	3	US-09-949-016-4758	Sequence 4758, Ap
21	302	8.2	1983	3	US-09-949-016-2066	Sequence 2066, Ap
22	302	8.2	3089	3	US-09-949-016-641	Sequence 641, App
23	298	8.1	2123	3	US-09-194-468A-29	Sequence 29, Appl
24	294.5	8.0	567	2	US-08-142-449B-5	Sequence 5, Appli
25	225	6.1	11665	3	US-09-949-016-12446	Sequence 12446, A
26	225	6.1	11665	3	US-09-949-016-16500	Sequence 16500, A
27	217.5	5.9	4588	3	US-08-840-062-1	Sequence 1, Appli
28	208	5.6	4627	3	US-09-949-016-935	Sequence 935, App
29	208	5.6	5633	3	US-09-023-655-1490	Sequence 1490, Ap
30	199.5	5.4	4892	3	US-09-976-594-167	Sequence 167, App
31	199.5	5.4	5158	3	US-09-023-655-1347	Sequence 1347, Ap
32	197	5.3	4621	3	US-09-949-016-4577	Sequence 4577, Ap
33	197	5.3	5627	3	US-09-949-016-2444	Sequence 2444, Ap
34	196.5	5.3	29954	3	US-09-949-016-13808	Sequence 13808, A
35	196.5	5.3	31040	3	US-09-949-016-12383	Sequence 12383, A
36	193.5	5.2	4464	2	US-08-400-159-7	Sequence 7, Appli
37	193.5	5.2	4483	3	US-08-611-729A-7	Sequence 7, Appli
38	193.5	5.2	4483	3	US-09-195-524-7	Sequence 7, Appli
39	193.5	5.2	4483	3	US-09-310-685-5	Sequence 5, Appli
40	193	5.2	3955	3	US-09-214-278-4	Sequence 4, Appli
41	193	5.2	3955	3	US-09-855-722-4	Sequence 3, Appli
42	189.5	5.1	4771	3	US-08-840-062-3	Sequence 30388, A
43	189	5.1	601	3	US-09-949-016-30388	Sequence 30391, A
44	189	5.1	601	3	US-09-949-016-30391	Sequence 70162, A
45	189	5.1	601	3	US-09-949-016-70162	

#### ALIGNMENTS

##### RESULT 1

US-09-566-921-135  
; Sequence 135, Application US/09566921  
; Patent No. 6682888  
; GENERAL INFORMATION:  
; APPLICANT: Loring, Jeanne F.  
; APPLICANT: Tingley, Debora W.  
; APPLICANT: Edwards, Carla M.  
; TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE  
; FILE REFERENCE: PA-0024 US  
; CURRENT APPLICATION NUMBER: US/09/566,921  
; CURRENT FILING DATE: 2000-05-05  
; NUMBER OF SEQ ID NOS: 138  
; SOFTWARE: PERL Program  
; SEQ ID NO 135  
; LENGTH: 8044  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6682888 427813.14  
US-09-566-921-135

Alignment Scores:  
Pred. No.: 5 87e-312 Length: 8044  
Score: 3592.50 Matches: 629  
Percent Similarity: 97.2% Conservatives: 0  
Best Local Similarity: 97.2% Mismatches: 3  
Query Match: 97.4% Indels: 15  
DB: 3 Gaps: 1

US-09-581-651D-2 (1-642) x US-09-566-921-135 (1-8044)



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 12, 2006, 10:25:20 ; Search time 9261.74 Seconds  
(without alignments)  
4864.742 Million cell updates/sec

Title: US-09-581-651D-2  
Perfect score: 3687  
Sequence: 1 MLRCPGGLLLAVQCLGTA.....ISKYLWRPVSIPPRNLGY 642

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/abss/ABSSWEB spool/US09581651/runat 10052006 181209 176/app query.fasta\_1  
-DB=EST -OPWT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs03h  
-USER=US09581651 @CGN 1.1 6731 @runat 10052006 181209 176 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_est3.\*  
4: gb\_est4.\*  
5: gb\_est5.\*  
6: gb\_est6.\*  
7: gb\_est7.\*  
8: gb\_est8.\*  
9: gb\_est9.\*  
10: gb\_est10.\*  
11: gb\_est11.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Length	ID	Description
1	3592.5	97.4	7868	4 HSM806653 BX640608 Homo sapi
2	3592.5	97.4	8411	4 HSM806992 BX640875 Homo sapi
3	3591.5	97.4	7885	4 CR749281 CR749316 Homo sapi
4	3588.5	97.3	7777	4 CR749316 CR749317 Homo sapi
5	3567.5	96.8	8121	4 CR749317 CR749317 Homo sapi
6	3564.5	96.7	7501	4 BC078656 BC078656 Homo sapi
7	3564.5	96.7	7501	4 BC100030 BC100030 Homo sapi

8	3389.5	91.9	8315	4 AK090135
9	3278	88.9	8329	4 AK090130
10	2485.5	67.4	7434	11 DQ039102
11	1804.5	48.9	7434	11 DQ039103
12	1573.5	42.7	943	5 BX391752
13	1564	42.4	1044	5 BX398837
14	1496	40.6	2107	4 AK054456
15	1490.5	40.4	836	7 CN538822
16	1480	40.1	1052	5 BX380582
17	1478	40.1	816	1 AU141008
18	1464	39.7	859	6 CF616056
19	1460	39.6	766	7 CN419594
20	1451.5	39.4	861	1 AU140971
21	1419.5	38.5	785	3 CN534124
22	1415	38.4	761	3 BM715855
23	1400	38.0	808	7 CK638401
24	1390	37.7	837	1 AU140973
25	1386	37.6	731	1 AU140526
26	1381	37.5	737	1 AU140993
27	1369	37.1	959	5 BU109952
28	1356	36.8	739	1 AL706215
29	1356	36.8	1026	5 BX386270
30	1351	36.6	695	7 CN419477
31	1349	36.6	729	1 AU140834
32	1342	36.4	849	8 DR835543
33	1340	36.3	772	7 CN530412
34	1336	36.2	799	8 CX899584
35	1329.5	36.1	781	7 CN161442
36	1323	35.9	699	1 AU140889
37	1309.5	35.5	856	8 CX894723
38	1308.5	35.5	1050	5 BX417945
39	1308	35.5	723	7 CN419622
40	1304.5	35.4	780	7 CN162843
41	1303	35.3	738	6 CD613781
42	1302.5	35.3	748	1 AU140948
43	1299	35.2	829	8 CX901794
44	1298	35.2	667	1 AU140910
45	1295.5	35.1	796	1 AU140814

## ALIGNMENTS

## RESULT 1

HSM806653

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

CONSTRM

TITLE

JOURNAL

COMMENT

HSM806653 7868 bp mRNA linear HTC 20-JAN-2005  
Homo sapiens mRNA; cDNA DKFZp686M04163 (from clone DKFZp686M04163).  
BX640608  
BX640608.1 GI:34364616  
HTC.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1. (bases 1 to 7868)  
Bloeker, H., Boecker, M., Brandt, P., Mewes, H.W., Weil, B., Amid, C.,  
Osanger, A., Fobo, G., Han, M. and Wiemann, S.  
The German cDNA Consortium  
Direct Submission  
Submitted (20-JAN-2005) MIPS, Ingolstaedter Landstr.1, D-85764  
Neuherberg, GERMANY  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by GBF (National Research Centre for Biotechnology Ltd.,  
Braunschweig/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
This clone (DKFZp686M04163) is available at the RZPD Deutsches  
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.  
Please contact RZPD for ordering:  
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp686M04163  
Further information about the clone and the sequencing project is  
available at http://mips.gsf.de/projects/cdna/.  
Location/Qualifiers

source	1. .7868	Query Match: DB:	97.4% 4	Indels: Gaps:	15 1
	/organism="Homo sapiens"	US-09-581-651D-2 (1-642) x HSM06653 (1-7868)			
	/mol_type="mRNA"				
	/db_xref="RZPD:DKFZp686M04163Q"				
	/db_xref="taxon:9606"				
	/clone="DKFZp686M04163"				
	/tissue_type="uterus endothel, primary cell culture"				
	/tissue_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host DH10B; sites SfiIA + SfiIB"				
	/dev_stage="adult"				
	/note="fibronectin precursor"				
gene	1. .7868				
	/gene="DKFZp686M04163"				
CDS	267..7157				
	/gene="DKFZp686M04163"				
	/codon_start=1				
	/product="hypothetical protein"				
	/protein_id="CA545714.1"				
	/db_xref="GI:34364617"				
	/db_xref="GOA:Q6N0A6"				
	/db_xref="InterPro:IPR000083"				
	/db_xref="InterPro:IPR000562"				
	/db_xref="InterPro:IPR003086"				
	/db_xref="InterPro:IPR003961"				
	/db_xref="InterPro:IPR006209"				
	/db_xref="UniProt/TREMBL:Q6N0A6"				
	/translation="MLRGPGLLLALLVLCVLAVPSTGASKSKRAQAOVMQPOSPVA VSQSKPGCYDNGKHVOINOQWERTVILGNALVCTCYGSGRGFNCBSKPEABETCFDKYT GNTYRGVDYERPKDSMIWDCITCIGAGRGRISCTIANRCHGGQSYKIGDWRPHEHT GNYMLSCVLNGNGKSWTCPIAEKCFDHAAGTSYVGVETWEKPYQGMWMDVCTCLGE GSGRIITCRRNRCDQDTYSRIIGDTSKKDNCRNLQOCI CTGNGRGKWKCEKHSY QTTSQSGPFTDVAAYQOPHPQPPYPYCHVCITDGSVYVGMWLKIQGNKQMLCT CLNGVSCQETAIVOTYQYNGSNGEPCVLPFTYNGRTFYSCITTEGRQDGLHWCSTTSY EQDKYSFCETHVLTQYTGNSNGALCHFPFLYNNHNYTCTSEGRDNKWKGGITQ NYDADQSPFCPWAHEEICTNNEGRIQDQDKDQMDGMMRCCTCVNGRGWET IAYSQRDQICVDIITVYNDYDFHKGHEEGLHMLCTCFQGRGRWKCDPVDQCDST GTPTQIGSWEKYHGVRYQCYCYRGIGGEWHQCPLOTYPSSSGPVEVFIETPSPQN SHIPQWNAPOPSHISIKYILRWPKNSVGRWKEATIPGLNSVTYIKGLKPGVYVYEGOL SIOQYGHQVETDFDTTSTSTPVSNTVTGTTTTPSPVATSESVTEITASSPVFWS VSASDVSFGRVEYELSEGDSPQYLDLPSTATSVNI PDLPLPKRYIYVYVQISEDE QSILSTSQITPADPDPDQVDDTSILVWRMSRQAPITGTYRIVYSPVSGSSTSL NLPEFANSVTLSDLQPGVQNYITIIYAVEENQESTPVIIQETTGTTPRSDDVPSRDLQ FVEYTDVKTIVMTPESAVTGYRVDVIVPNLPGHGORLPISRNTEAEVTGLSPGYT YFKFPAVSHGRESKPLTAQOTTKLADPNLQFVNETDSTVLVRWTPPRAQITGYRLT VGLTRGQPTVENVYVSGVSKYPLRNLOPASEVYVSLVAIKGNQSPKATGYFTTLQPG SSTIPPYNTVEYTVITWTPAPRIGFKLGVRSQCGEAPREVTSDSGSIUVSGLTQPG VEYVYITQVLRQKQDAPLVNKPVTPLUSPPTNLHLEANPDGTGLTVSWERSTTPDIT GYRITFTPTNGQGSLSLEEVHADQSCSTFNLSPGLEYNVSVYTKDDKESVPIST IIPAVPPPTDLRAFTNGIDPMVTRWTPAPPSIDLNFLVYSPVKNESDVAELSIPSD NAVLNLPGTYEYVYVSVSVEQHESTPLRGOKTGLDSTPGIDFSDITANSSTVHM IAPRATITGVRIHHPHEHPSGRPRDVRPHSRNSITLNLTPGTEYVYVSVIALNGREE SPLLICQQTSDVSDPDLLEVAATPTSLIISWDAPAVTVRYRITYGETGNSVPQBF TVPGSKTSQITPADPDPDQVDDTSILVWRMSRQAPITGTYRIVYSPVSGSSTSL VQGNSSVWMLSPSSGPYGYRVTITPKNGPGPTKTKTAGPDQTEMTIEGLQPTVEYV SVYAPQSGESQPLVQTAVTITPAPTDLTKFTQVTPSLSAQMTSPNVLQGVTVRVRTP KEKTPGPKMGINLAPDSSVWVSGLVATKYEVSVYALKDITLTSRPAQGVTTLENVSP PRARVTDATETITLISWTKETITGQVDAVPANGQTPQRTIKPDVRSYITGLQ PGTDYKIYTLNDNARSPPVIDASTADLPSNLRFLATTPNSLLVSWQPPRARITG YIIEKPGSPPREVVRPVRPGTEATITGLEPTEYTYIVYIALKNKQKEPGLRKK TDELPLQVLTLPNHLGPEILDVSTVQKTPFTVHPGDVGTNGIQLQSPQSQSVQK QMTFESHGFRRTPTPTATPIRHRPRYPNNVGEI QIIGHIPREDVDYHLPHPGQLN PNASTQEAALSDTTISWAPFQDTSVYIISCHPVGTDEEPLQFRVPGYSTSATLTGLTR GATYNIIVZALQDQRHKVRSVTVGNVSEGLNQVPTDDSCFPDPTVYSHVAVGDEWE RMSESGFKLLCCQLGQSGFKRCDSSRWCHNGVNIQIKERKWRQGNQMSCTLIG NKGPEGKDPHEATCYDQKGTVHGEQWKEYLGAICSTCFQGGQRCDNCRPPGG EPTPEGTQGSYNOYSQRYHQRTNTNVNCPICEFMPDLVDQADREDSRE"				
ORIGIN					
Alignment Scores:					
Pred. No.:	0	Length:	7868		
Score:	3592.50	Matches:	629		
Percent Similarity:	97.2%	Conservative:	0		
Best Local Similarity:	97.2%	Mismatches:	3		



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 12, 2006, 09:52:22 ; Search time 1039.8 Seconds  
(without alignment)  
6172.414 Million cell updates/sec

Title: US-09-581-651D-2

Perfect score: 3687

Sequence: 1 MLRPGPGGULLLAQCGLTA.....ISKYILWRPVSIPRNLGY 642

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/abs/ABSSWEB.spool/US09581651/runat\_10052006\_181204\_141/app\_query.fasta 1  
-DB=N Geneseq -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs03p  
-USBR=US09581651 @CGN 1.1359 @runat\_10052006\_181204\_141 -NCPUP=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq 21.\*

1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002as.\*  
7: Geneseqn2002bs.\*  
8: Geneseqn2003as.\*  
9: Geneseqn2003bs.\*  
10: Geneseqn2003cs.\*  
11: Geneseqn2003ds.\*  
12: Geneseqn2004as.\*  
13: Geneseqn2004bs.\*  
14: Geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3687	100.0	2147	2 AAX81299	Aax81299 Human mig
2	3680	99.8	1929	13 ADS17488	Adsl17488 Nucleotid
3	3680	99.8	1929	13 ADR97657	Adr97657 Human fib
4	3680	99.8	1929	14 ADW44479	Adw44479 Human fib

#### ALIGNMENTS

RESULT 1

AAX81299

ID AAX81299 standard; cDNA; 2147 BP.

XX

AC AAX81299;

XX

DT 21-SBP-1999 (first entry)

XX

DE Human migration stimulating factor (MSF) 1-alpha encoding cDNA.

XX

KW Migration stimulatory factor; MSF; cell migration; modulation; human;

KW wound healing; scarring; MSF1-alpha; ss.

XX

OS Homo sapiens.

XX

FN WO9931233-A1.

XX

PD 24-JUN-1999.

XX

PF 15-DEC-1998; 98WO-GB003766.

XX

PR 16-DEC-1997; 97GB-00026539.

XX

PA (UYDU-) UNIV DUNDEE.

XX

PI Schor SL, Schor AM;

XX

DR WPI; 1999-430039/36.

Ady55703 Human fib  
Ady32457 Human fib  
Aeb56257 Event seq  
Aeb78061 Human DNA  
Adr67201 Human bla  
Adq38575 Human SNP  
Adl18477 Human pro  
Adz26738 Human fib  
Adq38581 Human SNP  
Adr66637 Human pro  
Adr65953 Human pro  
Acc06170 Human cDN  
Acc00412 Human cel  
Adp26740 Human fib  
Adp64998 Human fib  
Adg89565 Human fib  
Adq29601 Human col  
Adr67200 Human bla  
Adv70129 Tumor-ass  
Aea04397 Human cDN  
Aeb17605 Human fib  
Adz26744 Human fib  
Adz26748 Human fib  
Adz26564 Human fib  
Adw81010 Fibronect  
Adx69342 DNA encod  
Adz26746 Human fib  
Adq38578 Human SNP  
Adq38582 Human SNP  
Adq38583 Human SNP  
Adq38577 Human SNP  
Adq38586 Human SNP  
Adq38579 Human FN  
Aac89889 Human FN  
Adq38584 Human SNP  
Adq38574 Human SNP  
Adq38587 Human SNP  
Adq38576 Human SNP  
Adq38580 Human SNP  
Adj56196 Zebrafish

DR P-PSDB: AAY28901.

PT Proteins with cell migration stimulatory activity used in treating wound  
PT and preventing scarring.

PT and preventing scarring.

PS Example 1; Fig 1; 86pp; English.

CC The invention provides a human migration stimulatory factor (MSF)  
CC protein. Host cells containing a replicable vector comprising the MSF  
CC encoding nucleic acid can be used for the recombinant production of the  
CC protein. The polypeptide can be used for modulating cell migration,  
CC healing a wound and for preventing scarring. The present sequence  
CC represents the nucleotide sequence encoding a human MSF1alpha protein

SQ Sequence 2147 BP; 575 A; 512 C; 575 G; 485 T; 0 U; 0 Other;

**Alignment Scores:**

Pred. No.:	1.19e-242	Length:	2147
Score:	3687.00	Matches:	642
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	100.0%	Indels:	0
DB:	2	Gaps:	0

Score: 3687.00

Percent Similarity: 100.0%

Best Local Similarity: 100.0%

Query Match: 100.0%

DB: 2

US-09-581-651D-2 (1-642) x AAX81299 (1-2147)

1 MetLeuAraGlvproGlvproGlvLeuLeuLeuAlaValGlnCysIleuGlyThrAla 20

Db  
57 ATGCTTAGGGGTCGGGGCCGGGGTGTGGTGGCCGTCCAGTGCCTGGGACAGCG 116

21 ValProSerThrGlyValAserIvsSerIvsAraGlnAlaGlnGlnMetValGlnProGln 40

117 GTGCCCCCAAGGGAGCCTCGAAGAGCAAGAGCCAGGCTCAGCAATATGGTTACGCCCGC 176

0v 4] SerProVa]A]aVa]SerC]nSerI.vgProC]vCvST.vgandnC]vI.vgH;CT.vC]n 60

db 177 TCCCCCGCTGGCTCAGTCAGACGACGCCCGCTTGTATTCACCATATCCATTCAC

06.01.2019 10:45:11 [Umsatz] 12.19

237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 12

[illegible]

297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 12

101 E. 1st St. N. Minneapolis, Minn. 55401

[illegible]

C  
C  
C  
C  
E  
C  
C  
C  
F  
C  
C  
C  
T  
C  
C  
C  
L  
E  
C  
I  
E  
C  
C  
C

417

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

\_\_\_\_\_

\_\_\_\_\_

100

[illegible]

.....

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 12, 2006, 10:13:41 ; Search time 11858.3 Seconds  
(without alignments)  
4616.200 Million cell updates/sec

Title: US-09-581-651D-2

Perfect score: 3687

Sequence: 1 MLRGPGGLLLAVQCLGTA.....ISKYLWRVPSIPPNLGY 642

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/abs/ABSSWEB.spool/US09581651/runat.10052006.181206.153/app query fasta.1  
-DB=GenEmbl -OFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs04  
-USER=US09581651 @CGN 1.1.5142 @runat.10052006.181206.153 -NCPUs=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DISPATCH=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.\*  
1: gb.ba.\*  
2: gb.in.\*  
3: gb.env.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pr.\*  
9: gb.ro.\*  
10: gb.sts.\*  
11: gb.sy.\*  
12: gb.un.\*  
13: gb.vi.\*  
14: gb.htg.\*  
15: gb.pl.\*

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3687	100.0	2147	6	BD137021 Polypepti
2	3687	100.0	2147	6	AX003229 Sequence
3	3687	100.0	2147	8	AJ276395 Homo sapi

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

4	3680	99.8	1929	6	CQ871810	Sequence
5	3680	99.8	1929	6	CQ871828	Sequence
6	3680	99.8	2358	6	CQ875358	Sequence
7	3648.5	99.0	2192	8	HSA535086	Homo sapi
8	3648.5	99.0	5984	8	AJ849445	Homo sapi
9	3638.5	98.7	4321	8	HSM806214	Homo sapi
10	3599.5	97.6	7242	6	CQ896577	Sequence
11	3599.5	97.6	7753	8	AB191261	Homo sapi
12	3599.5	97.6	8027	6	CQ833991	Sequence
13	3599.5	97.6	8027	6	CQ875357	Sequence
14	3599.5	97.6	8815	6	CS023820	Sequence
15	3599.5	97.6	8815	6	CS037064	Sequence
16	3592.5	97.4	7502	8	HSM804082	Homo sapi
17	3592.5	97.4	7745	8	AB209840	Homo sapi
18	3592.5	97.4	7951	8	HSM806901	Homo sapi
19	3592.5	97.4	7951	8	HSM806902	Homo sapi
20	3592.5	97.4	8044	6	AR454662	Sequence
21	3592.5	97.4	8044	6	AX281712	Sequence
22	3589.5	97.4	8030	8	HSM806170	Homo sapi
23	3589.5	97.4	8035	8	HSM806171	Homo sapi
24	3588.5	97.3	8042	8	HSM806805	Homo sapi
25	3585.5	97.2	7544	8	HSM803509	Homo sapi
26	3580.5	97.1	8320	8	HSM806267	Homo sapi
27	3580.5	97.1	8421	8	HSM806903	Homo sapi
28	3578.5	97.1	7161	6	CS017033	Sequence
29	3578.5	97.1	7803	6	AR051657	Sequence
30	3448.5	93.5	7679	6	AR274901	Sequence
31	3448.5	93.5	7680	6	AR380744	Sequence
32	3448.5	93.5	7680	6	AX277596	Sequence
33	3448.5	93.5	7680	6	AX335368	Sequence
34	3448.5	93.5	7680	6	HSFIB1	Human mRNA
35	3433.5	93.1	7705	6	A14133	Fibronectin
36	3433.5	93.1	7705	6	AR034630	Sequence
37	3433.5	93.1	7705	6	E01162	cDNA encodi
38	3433.5	93.1	7705	6	I70110	Sequence 16
39	3428.5	92.0	7705	6	AR364992	Sequence
40	3393.5	93.0	3059	9	BC051082	Mus muscu
41	3354.5	91.0	8329	6	AX402055	Sequence
42	3354.5	91.0	8329	9	RNF1BTRON	Sequence
43	3289.5	89.2	6816	6	CQ715726	Rat mRNA fo
44	2897	78.6	8372	5	BC095906	Xenopus t
45	2888	78.3	8313	5	BC072841	Xenopus t

#### ALIGNMENTS

RESULT 1	BD137021	2147 bp	DNA	linear	PAT 18-SEP-2002
LOCUS	Polypeptides, polynucleotides and uses thereof.				
DEFINITION	Polypeptides, polynucleotides and uses thereof.				
ACCESSION	BD137021				
VERSION	BD137021.1	GI:23231966			
KEYWORDS	JP 2002508179-A/1.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;				
TITLE	Hominidae; Homo.				
JOURNAL	1 (bases 1 to 2147)				
COMMENT	Schor, S.L. and Schor, A.M.				
	Polypeptides, polynucleotides and uses thereof				
	Patent: JP 2002508179-A 1 19-MAR-2002;				
	UNIVERSITY OF DUNDEE				
	OS Homo sapiens (human)				
	PN JP 2002508179-A/1				
	PD 19-MAR-2002				
	PF 15-DEC-1998	JP 2000539133			
	PR 16-DEC-1997	GB 9726539.1			
	PI SETH LAWRENCE SCHOR, ANA MARIA SCHOR				
	PC C12N15/09, A61K38/00, A61P17/02, C07K14/78, C07K16/18, C12N5/10, PC				
	C12P21/09,				
	Polypeptides, polynucleotides and uses thereof. FH Key				

Location/Qualifiers		1. .2147		/organisms='Homo sapiens (human)'/	
FT	source	FT	source	FT	source
FEATURES		Location/Qualifiers		1. .2147	
source		/organism='Homo sapiens'		/mol_type='genomic DNA'	
ORIGIN		/db_xref='taxon:9606'			
Alignment Scores:					
Pred. No.:	3.5e-247	Length:	2147		
Score:	3687.00	Matches:	642		
Percent Similarity:	100.0%	Conservative:	0		
Best Local Similarity:	100.0%	Mismatches:	0		
Query Match:	100.0%	Indels:	0		
DB:	6	Gaps:	0		
US-09-581-651D-2 (1-642) x BD137021 (1-2147)					
QY	1	MetLeuArgGlyProGlyProGlyLeuLeuLeuAlaValGlnCysLeuGlyThrAla	20		837
DB	57	ATGCTTAGGGTCCGGGCCCCGGCTGCTGCTGCTGCGCTCCAGTGCCTGGGACAGCG	116		281
QY	21	ValProSerThrGlyAlaSerLysSerLysArgGlnAlaGlnMetValGlnProGln	40		897
DB	117	GTCCCTCCACGGAGCCTCGAAGACCAAGAGGAGGAGGCTCAGCAATGGTTTCAGCCCCAG	176		957
QY	41	SerProValAlaValSerGlnSerLysProGlyCysTyrAspAsnGlyLysHisTyrGln	60		321
DB	177	TCCCCGGTGGCTGTCAAGTCAAGCAAGCCGGTGTGTATGACAAATGGAAACACTATCAG	236		1017
QY	61	IleAsnGlnGlnTrpGluArgThrTyrLeuGlyAsnAlaLeuValCysThrCysTyrGly	80		341
DB	237	ATAAATCAACAGTGGGAGCGGACCTACCTAGGCAATGCGTTGGTTGTACTTGTATGGA	296		1077
QY	81	GlySerArgGlyPheAsnCysGluSerLysProGluAlaGluGluThrCysPheAspLys	100		1136
DB	297	GGAAGCCGAGGTTTTAACTGCAGAGAGTAAACCTGAAGCTGAAGAGACTTGTCTTTGACAAG	356		1196
QY	101	TyrThrGlyAsnThrTyrArgValGlyAspThrTyrGluArgProLysAspSerMetIle	120		1256
DB	357	TACACTGGGAAACACTTACCAGTGGGTGACACTTATGAGCGTCTCTAAAGACTCCATGATC	416		1316
QY	121	TrpAspCysThrCysIleGlyAlaGlyArgGlyArgIleSerCysThrIleAlaAsnArg	140		1376
DB	417	TGGGACTGTACCTCGGGCTGGGGAGGAGGAGTAAGCTGTACCATCGCAACCCGC	476		1436
QY	141	CysHisGluGlyGlyGlnSerTyrLysIleGlyAspThrTrpArgArgProHisGluThr	160		1496
DB	477	TGCCATGAAGGGGTCCAGTCTACAAGATTGGTGACACCTGGAGGAGACCATGAGACT	536		1556
QY	161	GlyGlyTyrMetLeuGluCysValCysLeuGlyAsnGlyLysGlyGluTrpThrCysLys	180		1616
DB	537	GGTGGTTACATGTTAGAGTGTGTGCTCTGGTAATGGAAAGGAGAGATGGACCTGCAAG	596		1676
QY	181	ProIleAlaGluLysCysPheAspHisAlaAlaGlyThrSerTyrValValGlyGluThr	200		1736
DB	597	CCCATAGCTGAGAAGTGTGTTGATCATGCTCTGGGACTTCTATGTTGGTGGGAGAAACG	656		1796
QY	201	TrpGluLysProTyrGlnGlyTrpMetMetValAspCysThrCysLeuGlyGlySer	220		1856
DB	657	TGGGAGAGCCCTACCAAGGCTGGATGATGGTAGATTGTACTTGTCTGGGAGAGGCAGC	716		1916
QY	221	GlyArgIleThrCysThrSerArgAsnArgCysAsnAspGlnAspThrArgThrSerTyr	240		1976
DB	717	GGACGCATCACTTGCACTTCTAGAAATAGATGCAACGATCAGGACACAAAGGACATCCTAT	776		2036
QY	241	ArgIleGlyAspThrTrpSerLysLysAspAsnArgGlyAsnLeuGlnCysIleCys	260		2096
DB	777	AGAAATGGAGACACCTGGAGCAAGAGGATATTCAGGAAACCTGCTCCAGTGCATCTGC	836		2156
QY	261	ThrGlyAsnGlyArgGlyGluTrpLysCysGluArgHisThrSerValGlnThrThrSer	280		2216

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2006, 19:42:45 ; Search time 38.4018 Seconds  
(without alignments)  
784.888 Million cell updates/sec

Title: US-09-581-651D-2  
Perfect score: 3687  
Sequence: 1 MLRGPGLLLAVQLCTA.....ISKYLWRPVSIPRNLGY 642

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 250354 seqs, 4694837 residues

Total number of hits satisfying chosen parameters: 250354

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA New:

- 1: /SIDSS/ptodata/2/pubpaa/US08\_NEW\_PUB.pep1.\*
- 2: /SIDSS/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 3: /SIDSS/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 4: /SIDSS/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 5: /SIDSS/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 6: /SIDSS/ptodata/2/pubpaa/US05\_NEW\_PUB.pep.\*
- 7: /SIDSS/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 8: /SIDSS/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 9: /SIDSS/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 10: /SIDSS/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 11: /SIDSS/ptodata/2/pubpaa/US11\_NEW\_PUB.pep1.\*
- 12: /SIDSS/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3664	99.4	642	9	US-10-995-561-631
2	3653.5	99.1	657	11	US-11-193-561-27
3	3653.5	99.1	657	11	US-11-193-771-27
4	3653.5	99.1	657	11	US-11-193-789-27
5	3653.5	99.1	657	11	US-11-193-806-27
6	3653.5	99.1	657	11	US-11-193-857-27
7	3646.5	98.9	657	9	US-10-995-561-622
8	3599.5	97.6	2176	11	US-11-193-561-25
9	3599.5	97.6	2176	11	US-11-193-771-25
10	3599.5	97.6	2176	11	US-11-193-789-25
11	3599.5	97.6	2176	11	US-11-193-806-25
12	3599.5	97.6	2176	11	US-11-193-857-25
13	3599.5	97.6	2217	11	US-11-193-561-38
14	3599.5	97.6	2217	11	US-11-193-771-38
15	3599.5	97.6	2217	11	US-11-193-789-38
16	3599.5	97.6	2217	11	US-11-193-806-38
17	3599.5	97.6	2217	11	US-11-193-857-38
18	3599.5	97.6	2296	11	US-11-193-561-23
19	3599.5	97.6	2296	11	US-11-193-771-23
20	3599.5	97.6	2296	11	US-11-193-789-23
21	3599.5	97.6	2296	11	US-11-193-806-23

22	3599.5	97.6	2296	11	US-11-193-857-23	Sequence 23, Appl
23	3599.5	97.6	2330	11	US-11-193-561-21	Sequence 21, Appl
24	3599.5	97.6	2330	11	US-11-193-771-21	Sequence 21, Appl
25	3599.5	97.6	2330	11	US-11-193-789-21	Sequence 21, Appl
26	3599.5	97.6	2330	11	US-11-193-806-21	Sequence 21, Appl
27	3599.5	97.6	2330	11	US-11-193-857-21	Sequence 21, Appl
28	3599.5	97.6	2355	11	US-11-193-561-19	Sequence 19, Appl
29	3599.5	97.6	2355	11	US-11-193-771-19	Sequence 19, Appl
30	3599.5	97.6	2355	11	US-11-193-789-19	Sequence 19, Appl
31	3599.5	97.6	2355	11	US-11-193-806-19	Sequence 19, Appl
32	3599.5	97.6	2355	11	US-11-193-857-19	Sequence 19, Appl
33	3599.5	97.6	2384	9	US-10-821-234-1545	Sequence 1545, Ap
34	3599.5	97.6	2421	11	US-11-193-561-17	Sequence 17, Appl
35	3599.5	97.6	2421	11	US-11-193-771-17	Sequence 17, Appl
36	3599.5	97.6	2421	11	US-11-193-789-17	Sequence 17, Appl
37	3599.5	97.6	2421	11	US-11-193-806-17	Sequence 17, Appl
38	3599.5	97.6	2421	11	US-11-193-857-17	Sequence 17, Appl
39	3599.5	97.6	2477	11	US-11-193-561-15	Sequence 15, Appl
40	3599.5	97.6	2477	11	US-11-193-771-15	Sequence 15, Appl
41	3599.5	97.6	2477	11	US-11-193-789-15	Sequence 15, Appl
42	3599.5	97.6	2477	11	US-11-193-806-15	Sequence 15, Appl
43	3599.5	97.6	2477	11	US-11-193-857-15	Sequence 15, Appl
44	3592.5	97.4	984	9	US-10-995-561-629	Sequence 629, App
45	3592.5	97.4	2296	9	US-10-995-561-633	Sequence 633, App

## ALIGNMENTS

RESULT 1  
US-10-995-561-631  
; Sequence 631, Application US/10995561  
; Publication No. US2005027054A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
; FILE REFERENCE: CL001559  
; CURRENT APPLICATION NUMBER: US/10/995,561  
; CURRENT FILING DATE: 2004-11-24  
; NUMBER OF SEQ ID NOS: 85702  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 631  
; LENGTH: 642  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-995-561-631

Query Match	99.4%	Score 3664;	DB 9;	Length 642;
Best Local Similarity	99.5%	Pred. No. 9.7e-285;		
Matches 639;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
Qy	1	MLRGPGLLLAVQLCTA	VPSTGASKRQ	AAQMVQPQSPVAVSQKPGCYDNGHYQ 60
Db	1	MLRGPGLLLAVQLCTA	VPSTGASKRQ	AAQMVQPQSPVAVSQKPGCYDNGHYQ 60
Qy	61	INQWERTYLGNA	VCTCYGSGRGNCE	SKPEAETCFDKYTGNTYRVGTYRPPKDSMI 120
Db	61	INQWERTYLGNA	VCTCYGSGRGNCE	SKPEAETCFDKYTGNTYRVGTYRPPKDSMI 120
Qy	121	WDCCTCIGAGRG	ISCTIANRCHGGG	QSYKIGDTRRPHETGGYMLCVCILGNKGWETCK 180
Db	121	WDCCTCIGAGRG	ISCTIANRCHGGG	QSYKIGDTRRPHETGGYMLCVCILGNKGWETCK 180
Qy	181	PIAEKCFDHAAGT	SYVVGTEWKEPYQ	GMWVMDCTCLGEGSGRITCTSRNCNDQDRTSY 240
Db	181	PIAEKCFDHAAGT	SYVVGTEWKEPYQ	GMWVMDCTCLGEGSGRITCTSRNCNDQDRTSY 240
Qy	241	RIGDTWSKDN	RGNLLQCICTNGR	GEWKCERHTSVQTTSSSGSPPTDRAAVYQPPHP 300
Db	241	RIGDTWSKDN	RGNLLQCICTNGR	GEWKCERHTSVQTTSSSGSPPTDRAAVYQPPHP 300

QY 301 QPPYGHCVTDSDGVVYVSGWMLKTQGNKQMLCTCLNGVSCQETAVTQTYGNSNGEPC 360  
DB 301 QPPYGHCVTDSDGVVYVSGWMLKTQGNKQMLCTCLNGVSCQETAVTQTYGNSNGEPC 360  
QY 361 VLPFTYNDRTDSTTSNYEQDQKYSFCTDHTVLTQTRGNSNGALCHFPFLYNNHNTDCT 420  
DB 361 VLPFTYNGRTCSSTTSNYEQDQKYSFCTDHTVLTQTRGNSNGALCHFPFLYNNHNTDCT 420  
QY 421 SEGRDNMKCGTTONYDADQKFGFCPMAAHEEICTTNEGVMYRIGDQDKQHDGMHMR 480  
DB 421 SEGRDNMKCGTTONYDADQKFGFCPMAAHEEICTTNEGVMYRIGDQDKQHDGMHMR 480  
QY 481 CTCVNGRGWTCIAYSQLRDQCIYVDDITYNVNDTHKREHGHMLNCTCFGGGRGWKC 540  
DB 481 CTCVNGRGWTCIAYSQLRDQCIYVDDITYNVNDTHKREHGHMLNCTCFGGGRGWKC 540  
QY 541 DPVDCQDSETGTIFYQIGDSWEKYVHGVRQCYCYGRGIGEMHCQPLQTYPSSSGPVEVF 600  
DB 541 DPVDCQDSETGTIFYQIGDSWEKYVHGVRQCYCYGRGIGEMHCQPLQTYPSSSGPVEVF 600  
QY 601 ITETPSQPNHPIQWNAPOPSHISKYLLRWRPVSIPPRLGY 642  
DB 601 ITETPSQPNHPIQWNAPOPSHISKYLLRWRPVSIPPRLGY 642

## RESULT 2

US-11-193-561-27  
; Sequence 27, Application US/11193561  
; Publication No. US20060024757A1  
; GENERAL INFORMATION:  
; APPLICANT: Huesa, Robert  
; APPLICANT: Shorter, Simon  
; TITLE OF INVENTION: Detection of Oncofetal Fibronectin for Selection of Concepti  
; FILE REFERENCE: 17101-080001/831  
; CURRENT APPLICATION NUMBER: US/11/193,561  
; PRIOR FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: 60/592,823  
; PRIOR FILING DATE: 2004-07-30  
; PRIOR APPLICATION NUMBER: 60/592,803  
; PRIOR FILING DATE: 2004-07-30  
; PRIOR APPLICATION NUMBER: 60/592,825  
; PRIOR FILING DATE: 2004-07-30  
; PRIOR APPLICATION NUMBER: 60/592,804  
; PRIOR FILING DATE: 2004-07-30  
; PRIOR APPLICATION NUMBER: 60/592,824  
; PRIOR FILING DATE: 2004-07-30  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 27  
; LENGTH: 657  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: GenBank NM 054034  
; DATABASE ENTRY DATE: 2005-06-10  
US-11-193-561-27

Query Match 99.1%; Score 3653.5; DB 11; Length 657;  
Best Local Similarity 97.4%; Pred. No. 6.9e-284;  
Matches 640; Conservative 0; Mismatches 2; Indels 15; Gaps 1;  
QY 1 MLRGPGLLLAVQCLGTAVPSTGASKSRQAQQMVQPSPVAVSQSKPGCYDNGKHQY 60  
DB 1 MLRGPGLLLAVQCLGTAVPSTGASKSRQAQQMVQPSPVAVSQSKPGCYDNGKHQY 60  
QY 61 INQWERTYLGALVCTCYGSGRGFNCSKPEAEETCFDKYTGNTYRVGDTYERPKDSMI 120  
DB 61 INQWERTYLGALVCTCYGSGRGFNCSKPEAEETCFDKYTGNTYRVGDTYERPKDSMI 120  
QY 121 WDCCTCAGRGRIISCTIANRCHGGGQSYKIGDTPRRPHETGGVYMLBVCVLGNGKGWTC 180  
DB 121 WDCCTCAGRGRIISCTIANRCHGGGQSYKIGDTPRRPHETGGVYMLBVCVLGNGKGWTC 180

QY 181 PIAEKCDFDAAGTCTSYVVGETWEKPYQGMWMDCTCLGEGSGRITCTSRNRCNDODTRTSY 240  
DB 181 PIAEKCDFDAAGTCTSYVVGETWEKPYQGMWMDCTCLGEGSGRITCTSRNRCNDODTRTSY 240  
QY 241 RIGDTWSKDNKGNLLQCICTGNGRGEMKCRHTSVQTTSSGSGPFTDVRAAVYQPPHP 300  
DB 241 RIGDTWSKDNKGNLLQCICTGNGRGEMKCRHTSVQTTSSGSGPFTDVRAAVYQPPHP 300  
QY 301 QPPYGHCVTDSDGVVYVSGWMLKTQGNKQMLCTCLNGVSCQETAVTQTYGNSNGEPC 360  
DB 301 QPPYGHCVTDSDGVVYVSGWMLKTQGNKQMLCTCLNGVSCQETAVTQTYGNSNGEPC 360  
QY 361 VLPFTYNDRT-----DSTTSNYEQDQKYSFCTDHTVLTQTRGNSNGALC 405  
DB 361 VLPFTYNGRTFSYCTTEGRQDGHLMCSTTSNYEQDQKYSFCTDHTVLTQTRGNSNGALC 420  
QY 406 HFPFLYNNHNTDCTSEGRDNMKCGTTONYDADQKFGFCPMAAHEEICTTNEGVMYRI 465  
DB 421 HFPFLYNNHNTDCTSEGRDNMKCGTTONYDADQKFGFCPMAAHEEICTTNEGVMYRI 480  
QY 466 GDQWDKQHDGMHMRCTCVNGRGWTCIAYSQLRDQCIYVDDITYNVNDTHKREHGHM 525  
DB 481 GDQWDKQHDGMHMRCTCVNGRGWTCIAYSQLRDQCIYVDDITYNVNDTHKREHGHM 540  
QY 526 LNCTCFGGGRGWKCDPVDCQDSETGTIFYQIGDSWEKYVHGVRQCYCYGRGIGEMHCQ 585  
DB 541 LNCTCFGGGRGWKCDPVDCQDSETGTIFYQIGDSWEKYVHGVRQCYCYGRGIGEMHCQ 600  
QY 586 PLQTYPSSSGPVEVFITETPSQPNHPIQWNAPOPSHISKYLLRWRPVSIPPRLGY 642  
DB 601 PLQTYPSSSGPVEVFITETPSQPNHPIQWNAPOPSHISKYLLRWRPVSIPPRLGY 657

## RESULT 3

US-11-193-771-27  
; Sequence 27, Application US/11193771  
; Publication No. US20060024722A1  
; GENERAL INFORMATION:  
; APPLICANT: Fisher-Colbrie, Mark  
; APPLICANT: Hickok, Durlin  
; APPLICANT: LaPointe, Jerome P.  
; TITLE OF INVENTION: Samples for Detection of Oncofetal Fibronectin and uses thereof  
; FILE REFERENCE: 17101-027001/828  
; CURRENT APPLICATION NUMBER: US/11/193,771  
; CURRENT FILING DATE: 2005-07-29  
; PRIOR APPLICATION NUMBER: 60/592,823  
; PRIOR FILING DATE: 2004-07-30  
; PRIOR APPLICATION NUMBER: 60/592,803  
; PRIOR FILING DATE: 2004-07-30  
; PRIOR APPLICATION NUMBER: 60/592,825  
; PRIOR FILING DATE: 2004-07-30  
; PRIOR APPLICATION NUMBER: 60/592,804  
; PRIOR FILING DATE: 2004-07-30  
; PRIOR APPLICATION NUMBER: 60/592,824  
; PRIOR FILING DATE: 2004-07-30  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 27  
; LENGTH: 657  
; TYPE: PRT  
; ORGANISM: Homo Sapiens  
; PUBLICATION INFORMATION:  
; DATABASE ACCESSION NUMBER: GenBank NM 054034  
; DATABASE ENTRY DATE: 2005-06-10  
US-11-193-771-27

Query Match 99.1%; Score 3653.5; DB 11; Length 657;  
Best Local Similarity 97.4%; Pred. No. 6.9e-284;  
Matches 640; Conservative 0; Mismatches 2; Indels 15; Gaps 1;  
QY 1 MLRGPGLLLAVQCLGTAVPSTGASKSRQAQQMVQPSPVAVSQSKPGCYDNGKHQY 60  
DB 1 MLRGPGLLLAVQCLGTAVPSTGASKSRQAQQMVQPSPVAVSQSKPGCYDNGKHQY 60

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2006, 19:41:10 ; Search time 243.212 Seconds  
(without alignments)  
1102.934 Million cell updates/sec

Title: US-09-581-651D-2  
Perfect score: 3687  
Sequence: 1 MLRGPGLLLAVQCLGTA.....ISKYLWRPVSIPRNLGY 642

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA Main.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.psp.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.psp.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.psp.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.psp.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.psp.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.psp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3664	99.4	642	4	US-10-741-601-354
2	3664	99.4	642	5	US-10-741-600-1066
3	3646.5	98.9	657	4	US-10-741-601-359
4	3646.5	98.9	657	5	US-10-741-600-1072
5	3599.5	97.6	2220	4	US-10-236-392-4
6	3599.5	97.6	2355	4	US-10-144-194A-104
7	3599.5	97.6	2355	4	US-10-447-161-3
8	3599.5	97.6	2355	4	US-10-734-564-94
9	3599.5	97.6	2355	5	US-10-491-566-104
10	3599.5	97.6	2355	5	US-10-852-335A-147
11	3599.5	97.6	2355	5	US-10-287-436A-436
12	3599.5	97.6	2355	5	US-10-287-436A-1137
13	3599.5	97.6	2355	6	US-11-040-130-28
14	3599.5	97.6	2386	5	US-10-618-281-32
15	3595.5	97.5	2386	3	US-09-961-403-1
16	3595.5	97.5	2386	4	US-10-788-792-206
17	3595.5	97.5	2386	5	US-10-868-577A-59
18	3595.5	97.5	2386	5	US-10-485-758-4
19	3595.5	97.5	2386	5	US-10-485-758-9
20	3592.5	97.4	984	4	US-10-741-601-356
21	3592.5	97.4	984	5	US-10-741-600-1069
22	3592.5	97.4	2296	4	US-10-741-601-363
23	3592.5	97.4	2296	5	US-10-741-600-1075
24	3592.5	97.4	2355	4	US-10-741-601-357
25	3592.5	97.4	2355	4	US-10-741-601-366
26	3592.5	97.4	2355	5	US-10-741-600-1067
27	3592.5	97.4	2355	5	US-10-741-600-1078

28 3592.5 97.4 2386 4 US-10-741-601-360 Sequence 360, App  
29 3592.5 97.4 2386 5 US-10-741-600-1071 Sequence 1071, App  
30 3505.5 95.1 2474 5 US-10-450-763-52638 Sequence 52638, A  
31 3448.5 93.5 2328 4 US-10-171-311-64 Sequence 64, Appl  
32 3448.5 93.5 2328 4 US-10-236-031B-70 Sequence 70, Appl  
33 3448.5 93.5 2328 4 US-10-374-979-98 Sequence 98, Appl  
34 3448.5 93.5 2328 4 US-10-182-936A-98 Sequence 677, App  
35 3448.5 93.5 2328 5 US-10-477-238A-677 Sequence 98, Appl  
36 3448.5 93.5 2328 5 US-10-680-287A-677 Sequence 677, App  
37 3448.5 93.5 2328 5 US-10-477-173-677 Sequence 677, App  
38 3439.5 93.3 2355 4 US-10-360-101-235 Sequence 235, App  
39 3424.5 92.9 2320 4 US-10-279-733-8 Sequence 8, Appl  
40 3410.5 92.5 2320 4 US-10-236-392-2 Sequence 2, Appl  
41 3296.5 89.4 1173 5 US-10-450-763-52634 Sequence 52634, A  
42 2141 58.1 463 4 US-10-144-194A-52 Sequence 52, Appl  
43 2141 58.1 463 5 US-10-491-566-52 Sequence 52, Appl  
44 1859.5 50.4 343 3 US-09-934-706-1 Sequence 1, Appl  
45 1859.5 50.4 400 3 US-09-934-706-5 Sequence 5, Appl

## ALIGNMENTS

RESULT 1  
US-10-741-601-354  
; Sequence 354, Application US/10741601  
; Publication No. US20040166519A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01500  
; CURRENT APPLICATION NUMBER: US/10/741,601  
; CURRENT FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 26415  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 354  
; LENGTH: 642  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-741-601-354

Query Match 99.4%; Score 3664; DB 4; Length 642;  
Best Local Similarity 99.5%; Pred. No. 1.2e-289;  
Matches 639; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MLRGPGLLLAVQCLGTAVPSTGASKSRQQAQMVQSPVAVSQSKPCYDNGKHQY 60  
Db 1 MLRGPGLLLAVQCLGTAVPSTGASKSRQQAQMVQSPVAVSQSKPCYDNGKHQY 60  
Qy 61 INQWERTYLGNAVCTCYGSGRGFNCESKPEAEETCFDKYTGTVRVGDTYERPKDSMI 120  
Db 61 INQWERTYLGNAVCTCYGSGRGFNCESKPEAEETCFDKYTGTVRVGDTYERPKDSMI 120  
Qy 121 WDCTCIGAGRGRICTIANRCHGGQSYKIGDTRWRPHETGGYMLCEVCLNGKGWTK 180  
Db 121 WDCTCIGAGRGRICTIANRCHGGQSYKIGDTRWRPHETGGYMLCEVCLNGKGWTK 180  
Qy 181 PIAEKCFDHAAGTSYVVGTEWXPYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 240  
Db 181 PIAEKCFDHAAGTSYVVGTEWXPYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 240  
Qy 241 RIGDTSWKDNRNLLQCICTGNRGKWCERHSTVQTTSSGSGPFTDVRAAVYQPPHP 300  
Db 241 RIGDTSWKDNRNLLQCICTGNRGKWCERHSTVQTTSSGSGPFTDVRAAVYQPPHP 300  
Qy 301 QPPYPCHVTDSDGVVYVSGVQWLKTQGNQMLCTCLNGVSCQETAVTQTYGNSNGEPC 360  
Db 301 QPPYPCHVTDSDGVVYVSGVQWLKTQGNQMLCTCLNGVSCQETAVTQTYGNSNGEPC 360  
Qy 361 VLPFTYNDRTSDTSNYEQDKYSFCTDHTVLVQTRGNSNGALCHFPFLYNNHNTDCT 420  
Db 361 VLPFTYNGRTCTSTTSNYEQDKYSFCTDHTVLVQTRGNSNGALCHFPFLYNNHNTDCT 420

```
QY 421 SEGRDNNKWCCTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRIGDQDKQHDGMHMR 480
Db 421 SEGRDNNKWCCTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRIGDQDKQHDGMHMR 480
QY 481 CTCVNGRGWETCIAYSQLRDQCIYDQKFGFCPMAAHEEICTTNEGVMYRIGDQDKQHDGMHMR 540
Db 481 CTCVNGRGWETCIAYSQLRDQCIYDQKFGFCPMAAHEEICTTNEGVMYRIGDQDKQHDGMHMR 540
QY 541 DPVDCODSETGTFFQIGDSWEKYVHGVRVQCYCYGRGIGEWHCQPLQTYPSSSGPVEVF 600
Db 541 DPVDCODSETGTFFQIGDSWEKYVHGVRVQCYCYGRGIGEWHCQPLQTYPSSSGPVEVF 600
QY 601 ITETSPQNSHPQIOWNAPOQSHISKYILRWRPVSIPPRNLGY 642
Db 601 ITETSPQNSHPQIOWNAPOQSHISKYILRWRPVSIPPRNLGY 642
RESULT 2
US-10-741-600-1066
; Sequence 1066, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1066
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1066
Query Match 99.4%; Score 3664; DB 5; Length 642;
Best Local Similarity 99.5%; Pred. No. 1.2e-289;
Matches 639; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 MLRGPGLLLAVCLGTAVPSTGASKSKRAQQMVQPQSPVAVSQSKPGCYDNGKHQ 60
Db 1 MLRGPGLLLAVCLGTAVPSTGASKSKRAQQMVQPQSPVAVSQSKPGCYDNGKHQ 60
QY 61 INQWERTYLGNAVLVCTCYGSGRGNCSKPEABETCFDKYTGNTYRVGDTYVERPKDSMI 120
Db 61 INQWERTYLGNAVLVCTCYGSGRGNCSKPEABETCFDKYTGNTYRVGDTYVERPKDSMI 120
QY 121 WDCTCIGAGRGRIISCTIANRCHGGQSYKIGDTRWRPHETGGYMLBVCVCLGNGKGWTC 180
Db 121 WDCTCIGAGRGRIISCTIANRCHGGQSYKIGDTRWRPHETGGYMLBVCVCLGNGKGWTC 180
QY 181 PIAEKCFDHAAGTSYVVGTEWKPQGMWVVDCTCLGEGSGRITCTSRNRCNDQDTRTSY 240
Db 181 PIAEKCFDHAAGTSYVVGTEWKPQGMWVVDCTCLGEGSGRITCTSRNRCNDQDTRTSY 240
QY 241 PIAEKCFDHAAGTSYVVGTEWKPQGMWVVDCTCLGEGSGRITCTSRNRCNDQDTRTSY 300
Db 241 PIAEKCFDHAAGTSYVVGTEWKPQGMWVVDCTCLGEGSGRITCTSRNRCNDQDTRTSY 300
QY 301 QPPPYGHCVTDSGVVYVSGMQLKTOGNKQMLCTCLGNGVSCQETAVTQTYGNSNGEPC 360
Db 301 QPPPYGHCVTDSGVVYVSGMQLKTOGNKQMLCTCLGNGVSCQETAVTQTYGNSNGEPC 360
QY 361 VLPFTYNDRTSTSYEQDQKYSFCTDHTVLVQTRGNSNGALC 405
Db 361 VLPFTYNDRTSTSYEQDQKYSFCTDHTVLVQTRGNSNGALC 405
QY 406 HPFFLYNNHNYTDCSTSEGRDNNKWCCTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRI 465
Db 406 HPFFLYNNHNYTDCSTSEGRDNNKWCCTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRI 465
QY 421 SEGRDNNKWCCTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRIGDQDKQHDGMHMR 480
Db 421 SEGRDNNKWCCTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRIGDQDKQHDGMHMR 480
```

```
QY 481 CTCVNGRGWETCIAYSQLRDQCIYDQKFGFCPMAAHEEICTTNEGVMYRIGDQDKQHDGMHMR 540
Db 481 CTCVNGRGWETCIAYSQLRDQCIYDQKFGFCPMAAHEEICTTNEGVMYRIGDQDKQHDGMHMR 540
QY 541 DPVDCODSETGTFFQIGDSWEKYVHGVRVQCYCYGRGIGEWHCQPLQTYPSSSGPVEVF 600
Db 541 DPVDCODSETGTFFQIGDSWEKYVHGVRVQCYCYGRGIGEWHCQPLQTYPSSSGPVEVF 600
QY 601 ITETSPQNSHPQIOWNAPOQSHISKYILRWRPVSIPPRNLGY 642
Db 601 ITETSPQNSHPQIOWNAPOQSHISKYILRWRPVSIPPRNLGY 642
RESULT 3
US-10-741-601-359
; Sequence 359, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 359
; LENGTH: 657
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-359
Query Match 98.9%; Score 3646.5; DB 4; Length 657;
Best Local Similarity 97.3%; Pred. No. 3.4e-288;
Matches 639; Conservative 0; Mismatches 3; Indels 15; Gaps 1;
QY 1 MLRGPGLLLAVCLGTAVPSTGASKSKRAQQMVQPQSPVAVSQSKPGCYDNGKHQ 60
Db 1 MLRGPGLLLAVCLGTAVPSTGASKSKRAQQMVQPQSPVAVSQSKPGCYDNGKHQ 60
QY 61 INQWERTYLGNAVLVCTCYGSGRGNCSKPEABETCFDKYTGNTYRVGDTYVERPKDSMI 120
Db 61 INQWERTYLGNAVLVCTCYGSGRGNCSKPEABETCFDKYTGNTYRVGDTYVERPKDSMI 120
QY 121 WDCTCIGAGRGRIISCTIANRCHGGQSYKIGDTRWRPHETGGYMLBVCVCLGNGKGWTC 180
Db 121 WDCTCIGAGRGRIISCTIANRCHGGQSYKIGDTRWRPHETGGYMLBVCVCLGNGKGWTC 180
QY 181 PIAEKCFDHAAGTSYVVGTEWKPQGMWVVDCTCLGEGSGRITCTSRNRCNDQDTRTSY 240
Db 181 PIAEKCFDHAAGTSYVVGTEWKPQGMWVVDCTCLGEGSGRITCTSRNRCNDQDTRTSY 240
QY 241 RIGDTSWKDNRNLLQCICTGNGRGWKCERHTSVQTTSSGSGPFTDVRAAVYQPPHP 300
Db 241 RIGDTSWKDNRNLLQCICTGNGRGWKCERHTSVQTTSSGSGPFTDVRAAVYQPPHP 300
QY 301 QPPPYGHCVTDSGVVYVSGMQLKTOGNKQMLCTCLGNGVSCQETAVTQTYGNSNGEPC 360
Db 301 QPPPYGHCVTDSGVVYVSGMQLKTOGNKQMLCTCLGNGVSCQETAVTQTYGNSNGEPC 360
QY 361 VLPFTYNDRTSTSYEQDQKYSFCTDHTVLVQTRGNSNGALC 405
Db 361 VLPFTYNDRTSTSYEQDQKYSFCTDHTVLVQTRGNSNGALC 405
QY 406 HPFFLYNNHNYTDCSTSEGRDNNKWCCTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRI 465
Db 406 HPFFLYNNHNYTDCSTSEGRDNNKWCCTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRI 465
QY 421 SEGRDNNKWCCTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRIGDQDKQHDGMHMR 480
Db 421 SEGRDNNKWCCTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRIGDQDKQHDGMHMR 480
QY 526 LNCTCFGQGRGWKCDPVDQDSETGTGYQTGDSWEKYVHGVRVQCYCYGRGIGEWHCQ 585
Db 526 LNCTCFGQGRGWKCDPVDQDSETGTGTGYQTGDSWEKYVHGVRVQCYCYGRGIGEWHCQ 585
```



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2006, 19:13:45 ; Search time 303.276 Seconds  
(without alignments)  
1493.521 Million cell updates/sec

Title: US-09-581-651D-2  
Perfect score: 3687  
Sequence: 1 MLRGPGLLLAVQCLGTA.....ISKYLWRPVSIPRNLGY 642

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt 05.80.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3648.5	99.0	657	2	Q564H7 HUMAN
2	3638.5	98.7	749	2	Q72391 HUMAN
3	3599.5	97.6	2217	2	Q585T2 HOMO SAPIEN
4	3599.5	97.6	2265	2	Q60FE4 HOMO SAPIEN
5	3595.5	97.5	2386	1	F1NC HUMAN
6	3592.5	97.4	1103	2	Q6MZF4 HUMAN
7	3592.5	97.4	2296	2	Q6N0A6 HUMAN
8	3592.5	97.4	2351	2	Q59EH1 HUMAN
9	3592.5	97.4	2477	2	Q6MZU5 HOMO SAPIEN
10	3591.5	97.4	2357	2	Q68DT4 HUMAN
11	3588.5	97.3	2267	2	Q68DP9 HUMAN
12	3588.5	97.3	2444	2	Q6N025 HOMO SAPIEN
13	3567.5	96.8	2240	2	Q68DP8 HUMAN
14	3393.5	92.0	2477	1	F1NC MOUSE
15	3354.5	91.0	2477	1	F1NC RAT
16	3328.5	90.3	2265	1	F1NC BOVIN
17	2897	78.6	2475	2	Q501R6 XENTR
18	2888	78.3	2481	2	Q6GQAS XENLA
19	2878	78.1	2481	1	F1NC XENLA
20	2513.5	68.2	922	2	Q93405 BRARE
21	2513.5	68.2	2478	2	Q93406 BRARE
22	2383	64.6	1395	2	Q4SHU0 TETNG
23	2328	63.1	2408	2	Q6JAN2 BRARE
24	2314	62.8	2500	2	Q58XP5 BRARE
25	2301.5	59.7	2383	2	Q4RNC4 TETNG
26	1942.5	52.7	360	2	Q5RLI3 PIG
27	1909.5	51.8	2193	2	Q6MZM7 HUMAN
28	1817	49.3	383	2	Q4KL80 MOUSE
29	1496	40.6	296	2	Q8C6I7 MOUSE
30	1303	35.3	231	2	Q5MD86 MACMU
31	961	26.1	190	1	F1NC NOTVI

32	750	20.3	141	2	Q90XQ2 AMBME	Q90xq2 ambystoma m
33	541	14.7	215	2	Q6DD34 XENLA	Q6dd34 xenopus lae
34	331	9.0	410	2	Q4RVU5 TETNG	Q4rvu5 tetraodon n
35	319	8.7	677	2	Q4JF83 FUGRU	Q4jf83 fugu rubrip
36	317.5	8.6	679	2	Q98856 CVNPY	Q98856 cynops pyrr
37	313.5	8.5	673	2	Q90YB3 PAROL	Q90yb3 paralichth
38	310.5	8.4	690	2	Q9PVM5 ORYLA	Q9pvm5 oryzias lat
39	307.5	8.3	679	2	Q5D7I3 NOTVI	Q5d7i3 notophthalm
40	307	8.3	670	2	Q6DFI6 XENTR	Q6dfi6 xenopus tro
41	307	8.3	680	2	Q7T3I7 BRARE	Q7t3i7 brachydanio
42	306	8.3	675	2	Q8OFO6 ONCMY	Q8ofq6 oncorhynch
43	302.5	8.2	810	2	Q8R3F3 MOUSE	Q8r3f3 mus musculu
44	302	8.2	671	2	Q9W7L6 XENLA	Q9w7l6 xenopus lae
45	302	8.2	671	2	Q6PF33 XENLA	Q6pf33 xenopus lae

ALIGNMENTS

RESULT 1

ID	Q564H7 HUMAN PRELIMINARY;	PRT;	657 AA.
AC	Q564H7;		
DC	10-MAY-2005 (TrEMBLrel. 30, Created)		
DT	10-MAY-2005 (TrEMBLrel. 30, Last sequence update)		
DE	10-MAY-2005 (TrEMBLrel. 30, Last annotation update)		
DE	Migration stimulating factor.		
OS	Name=FNI;		
GN	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;		
OC	Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RA	Kay R.A., Ellis I.R., Jones S.J., Perrier S., Florence M.M.,		
RA	Schor A.M., Schor S.L.;		
RT	"mRNA 3'-UTR truncation: a novel mechanism regulating gene		
RT	expression."		
RL	Submitted (OCT-2004) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AJ849445; CAH60958.1; -;		
DR	SMR; Q564H7; 48-140, 183-275, 305-464.		
DR	Ensembl; ENSG00000115414; Homo sapiens.		
DR	GO; GO:0005576; C:extracellular region; IEA.		
DR	InterPro; IPR006209; EGF-like.		
DR	InterPro; IPR000083; Fibrinctnl.		
DR	InterPro; IPR000562; FN type2 col bd.		
DR	InterPro; IPR001724; Glyco_hydro_58.		
DR	Pfam; PF00039; fn1; 9.		
DR	Pfam; PF00040; fn2; 2.		
DR	PRINTS; PR00012; FNTYPEI.		
DR	PRINTS; PR00013; FNTYPEII.		
DR	PRINTS; PR00849; GLHYDRLASE58.		
DR	SMART; SM00058; FNI; 9.		
DR	SMART; SM00059; FN2; 2.		
DR	PROSITE; PS00022; EGF_1; UNKNOWN_1.		
DR	PROSITE; PS01253; FNI_1; 9.		
DR	PROSITE; PS01091; FNI_2; 9.		
DR	PROSITE; PS00023; FN2_1; 2.		
DR	PROSITE; PS01092; FN2_2; 2.		
DR	Repeat.		
DR	SEQUENCE 657 AA; 73752 MW; 8CA04E64486ABCD0 CRC64;		
DR	Query Match	99.0%;	Score 3648.5; DB 2; Length 657;
DR	Best Local Similarity	97.3%;	Pred. No. 1.2e-263;
DR	Matches	639;	Conservative 0; Mismatches 3; Indels 15; Gaps 1;
Qy	1	MLRGPGLLLAVQCLGTAVPTGASKSKRQAQMVQPQSPVAVSQSKPGCYDNGKHQY	60
Db	1	MLRGPGLLLAVQCLGTAVPTGASKSKRQAQMVQPQSPVAVSQSKPGCYDNGKHQY	60
Qy	61	INQWERTYILGNALVCTCYGSGRGFNCSKPEAEETCFDKYTGTVRVGVTYERPKDSMI	120

```
Db 61 INQWERTYLGALVCTCYGSGRGFNCEKPEAETCFDKYTGNTYRVGDTYERPKDSMI 120
Qy 121 WDCTCIGAGRGRIISCTIANRCHGGQSYKIGDTWRRPHETGGYMLCVCCLGNKGKGEWTK 180
Db 121 WDCTCIGAGRGRIISCTIANRCHGGQSYKIGDTWRRPHETGGYMLCVCCLGNKGKGEWTK 180
Qy 181 PIAEKCFDHAAGTSYVVGETWEKPYQGMWVDCITCLGEGSGRITCTSRNRCNDQDTRTSY 240
Db 181 PIAEKCFDHAAGTSYVVGETWEKPYQGMWVDCITCLGEGSGRITCTSRNRCNDQDTRTSY 240
Qy 241 RIGDTWSKDNKRNLLQCICTGNRGKWKCEHRTSVQTTSSGSGPFTDVAAYVQPPHP 300
Db 241 RIGDTWRKKDNKRNLLQCICTGNRGKWKCEHRTSVQTTSSGSGPFTDVAAYVQPPHP 300
Qy 301 OPPPYGHCVTDGSGVYVSGHQLTKQGNKQMLCTCLGNVGSQETAVTQTYGSGNNEPC 360
Db 301 OPPPYGHCVTDGSGVYVSGHQLTKQGNKQMLCTCLGNVGSQETAVTQTYGSGNNEPC 360
Qy 361 VLPFTYNDRT-----DSTTSNYEQDQKYSFCTDHTVLVQTRGNSNGALC 405
Db 361 VLPFTYNGRTFYSCCTTEGRQDGLHWCSTTSNYEQDQKYSFCTDHTVLVQTRGNSNGALC 420
Qy 406 HFPFLYNNHNYTDCTSEGRDNKMKCGTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRI 465
Db 421 HFPFLYNNHNYTDCTSEGRDNKMKCGTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRI 480
Qy 466 GDQWDKQHDHGMHMRCTCVNGRGKWCETCIAYSQLRDQCIYVDDITYNVNDFTHKRHEEGH 525
Db 481 GDQWDKQHDHGMHMRCTCVNGRGKWCETCIAYSQLRDQCIYVDDITYNVNDFTHKRHEEGH 540
Qy 526 LNCTCFGQGRGWKCDPVDQDQSETGTFYQIGDSWEKYVHGVYOCYCYGRGIGSEWHCQ 585
Db 541 LNCTCFGQGRGWKCDPVDQDQSETGTFYQIGDSWEKYVHGVYOCYCYGRGIGSEWHCQ 600
Qy 586 PLQTYPSSSGPVEVFITETPSQNSHPIQWNAQPQSHISKYILRWPRVSIIPRNLCY 642
Db 601 PLQTYPSSSGPVEVFITETPSQNSHPIQWNAQPQSHISKYILRWPRVSIIPRNLCY 657

RESULT 2
Q72391 HUMAN
ID Q72391 HUMAN PRELIMINARY; PRT; 749 AA.
AC Q72391;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein DKFZp686B18150.
GN Name=DKFZp686B18150;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=human colon endothel primary cell culture;
RA Bloeker H., Boecker M., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX538045; CAD97984.1; -; mRNA.
DR HSSP; Q96KP7; 1FBR.
DR SMR; Q72391; 140-232, 275-367, 397-556.
DR GO; GO:0005576; C:extracellular region; IEA.
DR GO; GO:0008243; F:plasma membrane activator activity; IEA.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR000083; Fibronectin.
DR InterPro; IPR000562; FN Type II.
DR InterPro; IPR001724; Glyco_hydro_58.
DR Pfam; PF00039; fn1; 9.
DR Pfam; PF00040; fn2; 2.
DR PRINTS; PR00012; FNTYPEI.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00849; GLHYDRLASES8.
```

```
DR ProDom; PD000995; FN_Type_II; 2.
DR SMART; SM00058; FN1; 9.
DR SMART; SM00059; FN2; 2.
DR PROSITE; PS00022; EGF_1; UNKNOWN 1.
DR PROSITE; PS01253; FIBRONECTIN_1; 9.
DR PROSITE; PS00023; FIBRONECTIN_2; 2.
KW Hypothetical protein.
SQ SEQUENCE 749 AA; 83524 MW; C8DDF97F3ED2F0DE CRC64;

Query Match 98.7%; Score 3638.5; DB 2; Length 749;
Best Local Similarity 97.1%; Pred. No. 7.9e-263;
Matches 638; Conservative 0; Mismatches 4; Indels 15; Gaps 1;

Qy 1 MLRGPGPGLLLAVLCIGTAVPSTGASKRQAOQWVQSPVAVSQSPKGCYDNGKHQY 60
Db 93 MLRGPGPGLLLAVLCIGTAVPSTGASKRQAOQWVQSPVAVSQSPKGCYDNGKHQY 152
Qy 61 INQWERTYLGALVCTCYGSGRGFNCEKPEAETCFDKYTGNTYRVGDTYERPKDSMI 120
Db 153 INQWERTYLGALVCTCYGSGRGFNCEKPEAETCFDKYTGNTYRVGDTYERPKDSMI 212
Qy 121 WDCTCIGAGRGRIISCTIANRCHGGQSYKIGDTWRRPHETGGYMLCVCCLGNKGKGEWTK 180
Db 213 WDCTCIGAGRGRIISCTIANRCHGGQSYKIGDTWRRPHETGGYMLCVCCLGNKGKGEWTK 272
Qy 181 PIAEKCFDHAAGTSYVVGETWEKPYQGMWVDCITCLGEGSGRITCTSRNRCNDQDTRTSY 240
Db 273 PIAEKCFDHAAGTSYVVGETWEKPYQGMWVDCITCLGEGSGRITCTSRNRCNDQDTRTSY 332
Qy 241 RIGDTWSKDNKRNLLQCICTGNRGKWKCEHRTSVQTTSSGSGPFTDVAAYVQPPHP 300
Db 333 RIGDTWSKDNKRNLLQCICTGNRGKWKCEHRTSVQTTSSGSGPFTDVAAYVQPPHP 392
Qy 301 OPPPYGHCVTDGSGVYVSGHQLTKQGNKQMLCTCLGNVGSQETAVTQTYGSGNNEPC 360
Db 393 OPPPYGHCVTDGSGVYVSGHQLTKQGNKQMLCTCLGNVGSQETAVTQTYGSGNNEPC 452
Qy 361 VLPFTYNDRT-----DSTTSNYEQDQKYSFCTDHTVLVQTRGNSNGALC 405
Db 453 VLPFTYNGRTFYSCCTTEGRQDGLHWCSTTSNYEQDQKYSFCTDHTVLVQTRGNSNGALC 512
Qy 406 HFPFLYNNHNYTDCTSEGRDNKMKCGTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRI 465
Db 513 HFPFLYNNHNYTDCTSEGRDNKMKCGTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRI 572
Qy 466 GDQWDKQHDHGMHMRCTCVNGRGKWCETCIAYSQLRDQCIYVDDITYNVNDFTHKRHEEGH 525
Db 573 GDQWDKQHDHGMHMRCTCVNGRGKWCETCIAYSQLRDQCIYVDDITYNVNDFTHKRHEEGH 632
Qy 526 LNCTCFGQGRGWKCDPVDQDQSETGTFYQIGDSWEKYVHGVYOCYCYGRGIGSEWHCQ 585
Db 633 LNCTCFGQGRGWKCDPVDQDQSETGTFYQIGDSWEKYVHGVYOCYCYGRGIGSEWHCQ 692
Qy 586 PLQTYPSSSGPVEVFITETPSQNSHPIQWNAQPQSHISKYILRWPRVSIIPRNLCY 642
Db 693 PLQTYPSSSGPVEVFITETPSQNSHPIQWNAQPQSHISKYILRWPRVSIIPRNLCY 749

RESULT 3
Q585T2 HUMAN
ID Q585T2 HUMAN PRELIMINARY; PRT; 2217 AA.
AC Q585T2;
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Hypothetical protein FN1 (Fragment).
GN Name=FN1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
```

GenCore version 5.1.8

Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2006, 19:20:36 ; Search time 49.2331 Seconds  
(without alignments)  
1254.667 Million cell updates/sec

Title: US-09-581-651D-2

Perfect score: 3687

Sequence: 1 MLRGPGLLLLAQCGLTA.....ISKYLRRPVPSPRNLGY 642

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 80.\*

1: PIR1.\*

2: PIR2.\*

3: PIR3.\*

4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3579.5	97.1	2386	1 FNHU	fibronectin precur
2	3354.5	91.0	2477	2 S14428	fibronectin precur
3	3328.5	90.3	2265	1 FNBO	fibronectin - bovi
4	2888	78.3	2481	2 A43908	fibronectin - Afri
5	961	26.1	190	2 I51279	fibronectin - east
6	302.5	8.2	708	2 JC4364	gelatinase B (EC 3
7	301	8.2	662	2 S70365	gelatinase A (EC 3
8	300	8.1	662	2 A42496	gelatinase A (EC 3
9	300	8.1	662	2 S34780	gelatinase A (EC 3
10	300	8.1	707	1 A53796	gelatinase B (EC 3
11	298	8.1	663	1 S46492	gelatinase A (EC 3
12	297	8.1	660	1 A28153	gelatinase A (EC 3
13	294.5	8.0	708	2 S62907	gelatinase B (EC 3
14	294.5	8.0	712	1 I46031	gelatinase B (EC 3
15	294.5	8.0	730	1 I52580	gelatinase B (EC 3
16	294.5	8.0	730	2 JC1456	gelatinase B (EC 3
17	285.5	7.7	707	1 A34458	gelatinase B (EC 3
18	228.5	6.2	1020	2 A29355	fibronectin - chic
19	208	5.6	1455	1 A48925	mannose receptor p
20	199.5	5.4	1456	1 A36563	mannose receptor p
21	199	5.4	5376	2 T42215	zonadhesin - mouse
22	190.5	5.2	1479	2 T42710	mannose receptor,
23	188	5.1	1458	1 A49707	phospholipase A2 r
24	185.5	5.0	1326	2 B56395	secretory phosphol
25	185.5	5.0	1465	2 A56395	secretory phosphol
26	185	5.0	1584	2 T22674	hypothetical prote
27	181.5	4.9	473	2 A56375	adhesive plaque pr
28	179.5	4.9	722	2 I48324	DELTA-like 1 - mou
29	179.5	4.9	1463	2 A53210	phospholipase A2 r

## ALIGNMENTS

## RESULT 1

FNHU

fibronectin precursor [validated] - human

N;Alternate names: fibronectin splice form ED-A

C;Species: Homo sapiens (man)

C;Date: 27-Nov-1985 #sequence revision 31-Mar-1993 #text change 31-Dec-2004

C;Accession: A26460; A26284; S03917; A24854; A24476; A91008; A93529; A21011; A90495; A22

R;Dean, D.C.; Bowlus, C.L.; Bourgeois, S.

Proc. Natl. Acad. Sci. U.S.A. 84, 1876-1880, 1987

A;Title: Cloning and analysis of the promoter region of the human fibronectin gene.

A;Reference number: A26460; MUID:87175578; PMID:3031656

A;Accession: A26460

A;Molecule type: DNA

A;Residues: 1-49 &lt;DEA&gt;

A;Cross-references: UNIPROT:P02751; UNIPROT:Q14327; UNIPARC:UPI000016A926; GB:M15801; NI

R;Oldberg, A.; Ruoslahti, E.

J. Biol. Chem. 261, 2113-2116, 1986

A;Title: Evolution of the fibronectin gene.

A;Reference number: A26284; MUID:86111901; PMID:3003095

A;Accession: A26284

A;Molecule type: DNA

A;Residues: 1447-1540 &lt;OLD&gt;

A;Cross-references: UNIPARC:UPI0000112E37; GB:M12549; NID:G182698

A;Note: the authors translated the codon TTC for residue 1494 as Glu

R;Paolella, G.; Henschliffe, C.; Sebastio, G.; Baralle, F.E.

Nucleic Acids Res. 16, 3545-3557, 1988

A;Title: Sequence analysis and in vivo expression show that alternative splicing of ED-B

A;Reference number: S00848; MUID:88233940; PMID:3375063

A;Accession: S03917

A;Molecule type: DNA

A;Residues: 1594-1767, 'V' 1769-1783 &lt;PAO&gt;

A;Cross-references: UNIPARC:UPI000017432B; EMBL:X07718; NID:G31402

A;Note: the authors translated the codon AAC for residue 1631 as Asp

R;Vibe-Pedersen, K.; Magnusson, S.; Baralle, F.E.

FEBS Lett. 207, 287-291, 1986

A;Title: Donor and acceptor splice signals within an exon of the human fibronectin gene:

A;Reference number: A24854; MUID:87030929; PMID:3770201

A;Accession: A24854

A;Molecule type: DNA

A;Residues: 1992-2147 &lt;VIB&gt;

A;Cross-references: UNIPARC:UPI000017432C; GB:X04530; NID:G31436

R;Gutman, A.; Yamada, K.M.; Kornblihtt, A.

FEBS Lett. 207, 145-148, 1986

A;Title: Human fibronectin is synthesized as a pre-propolypeptide.

A;Reference number: A24476; MUID:87030890; PMID:3770189

A;Accession: A24476

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-14, 'Q', 16-38 &lt;GUT&gt;

A;Cross-references: UNIPARC:UPI000017432D

R;Kornblihtt, A.R.; Umezawa, K.; Vibe-Pedersen, K.; Baralle, F.E.

EMBO J. 4, 1755-1759, 1985

notch protein - fr  
thrombospondin 1 p  
Xotch protein - Af  
notch-1 protein -  
phospholipase-A(2)  
notch protein homo  
C-Delta-1 - chicke  
crumbs protein - f  
hypothetical prote  
probable laminin a  
coagulation factor  
laminin alpha-1 ch  
insulin-like growt  
cation-independent  
transmembrane prot

A;Title: Primary structure of human fibronectin: differential splicing may generate at l  
A;Reference number: A91008; MUID:85284965; PMID:2992939  
A;Accession: A91008  
A;Status: nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 32-1344; 1346-2080; 2112-2386 <KOR>  
A;Cross-references: UNIPARC:UPI0000174332F; UNIPARC:UPI0000174332F; GB:X02761  
R;Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.  
Nucleic Acids Res. 12, 5853-5868, 1984  
A;Title: Human fibronectin: cell specific alternative mRNA splicing generates polypeptid  
A;Reference number: A93529; MUID:84272258; PMID:6462919  
A;Accession: A93529  
A;Molecule type: mRNA  
A;Residues: 973-2080; 2112-2386 <K02>  
A;Cross-references: UNIPARC:UPI0000174332F; UNIPARC:UPI0000174330; GB:X00739  
R;Oldberg, A.; Linney, E.; Ruoslahti, E.  
J. Biol. Chem. 258, 10193-10196, 1983  
A;Title: Molecular cloning and nucleotide sequence of a cDNA clone coding for the cell a  
A;Reference number: A21011; MUID:83290929; PMID:6688418  
A;Accession: A21011  
A;Molecule type: mRNA  
A;Residues: 1434-1537 <OL2>  
A;Cross-references: UNIPARC:UPI0000174331; GB:K00055; NID:g182680; PIDN:AAAS2459.1; PID:  
R;Bernard, M.P.; Kolbe, M.; Weil, D.; Chu, M.L.  
Biochemistry 24, 2698-2704, 1985  
A;Title: Human cellular fibronectin: comparison of the carboxyl-terminal portion with ra  
A;Reference number: A90495; MUID:85280409; PMID:2392573  
A;Accession: A90495  
A;Molecule type: mRNA  
A;Residues: 1594-2386 <BER>  
A;Cross-references: UNIPARC:UPI000016A928; GB:M10905; NID:g182696; PIDN:AAAS2462.1; PID:  
R;Umezawa, K.; Kornblitt, A.R.; Baralle, F.E.  
FEBS Lett. 186, 31-34, 1985  
A;Title: Isolation and characterization of cDNA clones for human liver fibronectin.  
A;Reference number: A22245; MUID:85231203; PMID:2989004  
A;Accession: A22245  
A;Molecule type: mRNA  
A;Residues: 1948-2067 <UME>  
A;Cross-references: UNIPARC:UPI000000046A; GB:M27589; NID:g182705; PIDN:AAAS2465.1; PID:  
A;Accession: B22245  
A;Molecule type: mRNA  
A;Residues: 1975-1991; 2017-2039 <UM2>  
A;Cross-references: UNIPARC:UPI0000174332; UNIPARC:UPI0000174333; GB:M27590  
R;Sekiguchi, K.; Klos, A.M.; Kurachi, K.; Yoshitake, S.; Hakomori, S.  
Biochemistry 25, 4936-4941, 1986  
A;Title: Human liver fibronectin complementary DNAs: identification of two different mes  
A;Reference number: I52394; MUID:87026578; PMID:3021206  
A;Accession: I65273  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1978-1990, 2016-2018, 'N' 2020-2081, 2113-2127 <SEK>  
A;Cross-references: UNIPARC:UPI000006E04C; GB:M14060; NID:g182701; PIDN:AAAS2464.1; PID:  
R;Kornblitt, A.R.; Vibe-Pedersen, K.; Baralle, F.E.  
Proc. Natl. Acad. Sci. U.S.A. 80, 3218-3222, 1983  
A;Title: Isolation and characterization of cDNA clones for human and bovine fibronectins  
A;Reference number: A21165; MUID:83221567; PMID:6304699  
A;Accession: A21165  
A;Molecule type: mRNA  
A;Residues: 2291-2386 <K03>  
A;Cross-references: UNIPARC:UPI0000174334; GB:K00799; NID:g182681; PIDN:AAAS2460.1; PID:  
R;Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.  
J. Biol. Chem. 258, 12670-12674, 1983  
A;Title: Primary structure of human plasma fibronectin.  
A;Reference number: A92398; MUID:84032463; PMID:6630202  
A;Accession: A92398  
A;Molecule type: protein  
A;Residues: 32-47, 'C', 49-51, 'S', 53-72, 'A', 74-290 <GAR1>  
A;Cross-references: UNIPARC:UPI0000174335  
R;Garcia-Pardo, A.; Gold, L.I.  
Arch. Biochem. Biophys. 304, 181-188, 1993  
A;Title: Further characterization of the binding of fibronectin to gelatin reveals the F  
A;Reference number: S34791; MUID:93312001; PMID:8323285  
A;Accession: S34791

A;Molecule type: protein  
A;Residues: 291-300; 551-560 <GAR2>  
A;Cross-references: UNIPARC:UPI0000174336; UNIPARC:UPI0000174337  
R;Griffin, C.A.; Calaycay, J.; Shively, J.E.; Smith, R.L.  
Thromb. Res. 43, 469-477, 1986  
A;Title: Two plasma fibronectin fragments with different gelatin-binding properties.  
A;Reference number: A60904; MUID:87019725; PMID:3532418  
A;Accession: A60904  
A;Molecule type: protein  
A;Residues: 293-301 <GRI>  
A;Cross-references: UNIPARC:UPI0000174338  
R;Calaycay, J.; Pande, H.; Lee, T.; Borsi, L.; Siri, A.; Shively, J.E.; Zardi, L.  
J. Biol. Chem. 260, 12136-12141, 1985  
A;Title: Primary structure of a DNA- and heparin-binding domain (domain III) in human pla  
A;Reference number: A23901; MUID:86008277; PMID:3900070  
A;Accession: A23901  
A;Molecule type: protein  
A;Residues: 616-677, 'Q', 679-703, 'PT', <CAL>  
A;Cross-references: UNIPARC:UPI0000174339  
R;Pierschbacher, M.D.; Ruoslahti, E.; Sundelin, J.; Lind, P.; Peterson, P.A.  
J. Biol. Chem. 257, 9593-9597, 1982  
A;Title: The cell attachment domain of fibronectin. Determination of the primary structur  
A;Reference number: A92386; MUID:82265604; PMID:7050098  
A;Accession: A92386  
A;Molecule type: protein  
A;Residues: 1441-1548 <PIE>  
A;Cross-references: UNIPARC:UPI0000141CD5  
A;Note: residues 1524-1527 are responsible for the cell-binding activity  
R;Garcia-Pardo, A.; Rostagno, A.; Frangione, B.  
Biochem. J. 241, 923-928, 1987  
A;Title: Primary structure of human plasma fibronectin. Characterization of a 38 kDa dom  
A;Reference number: A32517; MUID:87241275; PMID:3593230  
A;Accession: A32517  
A;Molecule type: protein  
A;Residues: 1589-1630, 'T', 1722-2058 <GAR3>  
A;Cross-references: UNIPARC:UPI000017433A  
R;Tressel, T.; McCarthy, J.B.; Calaycay, J.; Lee, T.D.; Legesse, K.; Shively, J.E.; Pande  
Biochem. J. 274, 731-738, 1991  
A;Title: Human plasma fibronectin. Demonstration of structural differences between the A  
A;Reference number: S14357; MUID:91190085; PMID:2012601  
A;Accession: S14357  
A;Molecule type: protein  
A;Residues: 1614-1630, 'T', 1722-2081, 2113-2244 <TRE>  
A;Cross-references: UNIPARC:UPI000017433B  
R;Garcia-Pardo, A.; Pearlstein, E.; Frangione, B.  
J. Biol. Chem. 260, 10320-10325, 1985  
A;Title: Primary structure of human plasma fibronectin. Characterization of a 31,000-dalt  
A;Reference number: A23891; MUID:85261459; PMID:4019516  
A;Accession: A23891  
A;Molecule type: protein  
A;Residues: 2071-2080; 2112-2356 <GAR4>  
A;Cross-references: UNIPARC:UPI000017433C; UNIPARC:UPI000017433D  
C;Comment: The extra domain and connecting strand 3 are subject to developmental and tiss  
C;Comment: The cellular and plasma fibronectins are high molecular weight glycoproteins,  
action, and transformation.  
C;Genetics:  
A;Gene: GDB.FN1  
A;Cross-references: GDB:119135; OMIM:135600  
A;Map position: 2q34-q34  
A;Introns: 49/3; 1266/1; 1357/1; 1447/1; 1487/1; 1541/1; 1631/1; 1721/1; 1991/1; 2145/1  
C;Superfamily: fibronectin type I repeat homology; fibronectin type II repeat homology; f  
C;Keywords: acute phase; alternative splicing; cell adhesion; collagen binding; duplicat  
F;1-26/Domain: signal sequence #status predicted <SIG>  
F;27-31/Domain: propeptide #status predicted <PRO>  
F;32-2386/Product: fibronectin #status experimental <MAT>  
F;52-272/Domain: fibrin and heparin binding <PHB>  
F;52-87/Domain: fibronectin type I repeat homology <IF1>  
F;97-135/Domain: fibronectin type I repeat homology <IF2>  
F;141-179/Domain: fibronectin type I repeat homology <IF3>  
F;186-225/Domain: fibronectin type I repeat homology <IF4>  
F;231-270/Domain: fibronectin type I repeat homology <IF5>  
F;308-608/Domain: collagen binding <CBR>  
F;308-342/Domain: fibronectin type I repeat homology <IF6>

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2006, 19:25:45 ; Search time 65.9724 Seconds  
(without alignments)  
804.545 Million cell updates/sec

Title: US-09-581-651D-2

Perfect score: 3687

Sequence: 1 MLRGPGLLLAVQCLGTA.....ISKYLWRPVSIPPRNLGY 642

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/iaa/5\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/6\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/H\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/RE\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3595.5	97.5	2386	2	US-09-961-403-1
2	3579.5	97.1	2386	1	US-09-016-366A-12
3	3578.5	97.1	2446	4	PCT-US93-12687-2
4	3578.5	97.1	2446	4	PCT-US93-12687-2
5	3439.5	93.3	2355	2	US-10-360-101-235
6	3433.5	93.1	2327	6	5455158-1
7	3422.5	92.8	2324	1	US-08-283-857-1
8	3422.5	92.8	2324	4	PCT-US95-09819-1
9	3406.5	92.4	2231	1	US-08-153-799-16
10	371	10.1	65	1	US-08-982-597A-19
11	371	10.1	65	2	US-09-136-218-19
12	298	8.1	663	2	US-09-194-468A-30
13	297	8.1	631	2	US-08-448-489-17
14	297	8.1	631	2	US-08-689-730-17
15	297	8.1	660	2	US-08-704-711A-18
16	297	8.1	660	2	US-09-521-220-18
17	297	8.1	660	2	US-09-391-104-19
18	297	8.1	660	2	US-09-917-254-89
19	297	8.1	660	2	US-09-949-016-6512
20	297	8.1	660	2	US-09-949-016-7937
21	297	8.1	660	2	US-10-153-185-14
22	295	8.0	422	1	US-08-836-854-12
23	294.5	8.0	188	1	US-08-142-449B-14
24	294.5	8.0	474	1	US-08-836-854-9
25	292.5	7.9	429	2	US-09-194-468A-45
26	290	7.9	48	1	US-08-982-597A-22
27	290	7.9	48	2	US-09-136-218-22

28	285.5	7.7	707	2	US-08-704-711A-19	Sequence 19, Appl
29	285.5	7.7	707	2	US-09-521-220-19	Sequence 19, Appl
30	285.5	7.7	707	2	US-09-391-104-20	Sequence 20, Appl
31	285.5	7.7	707	2	US-09-949-016-6575	Sequence 6575, Ap
32	285.5	7.7	708	2	US-08-448-489-16	Sequence 16, Appl
33	285.5	7.7	708	2	US-09-689-730-16	Sequence 16, Appl
34	285.5	7.7	713	2	US-09-949-016-10629	Sequence 10629, A
35	278	7.5	48	1	US-08-982-597A-20	Sequence 20, Appl
36	278	7.5	48	1	US-09-136-218-20	Sequence 20, Appl
37	258	7.0	43	1	US-08-982-597A-21	Sequence 21, Appl
38	258	7.0	43	2	US-09-136-218-21	Sequence 21, Appl
39	220	6.0	96	1	US-08-717-169-2	Sequence 2, Appl
40	220	6.0	96	2	US-09-228-901A-2	Sequence 2, Appl
41	219.5	6.0	60	1	US-08-982-597A-18	Sequence 18, Appl
42	219.5	6.0	60	2	US-09-136-218-18	Sequence 18, Appl
43	208	5.6	1455	2	US-08-840-062-5	Sequence 5, Appl
44	206	5.6	37	1	US-08-982-597A-17	Sequence 17, Appl
45	206	5.6	37	2	US-09-136-218-17	Sequence 17, Appl

#### ALIGNMENTS

RESULT 1  
US-09-961-403-1  
; Sequence 1, Application US/09961403  
; Patent No. 6780594  
; GENERAL INFORMATION:  
; APPLICANT: HE-STUMPP, HOLGER  
; APPLICANT: HAENDLER, BERNARD  
; APPLICANT: KRAETZSCHMAR, JOERN  
; APPLICANT: KREFT, BERTHOLT  
; APPLICANT: WINTERHAGER, ELKE  
; APPLICANT: REGIOR, PEDRO  
; APPLICANT: SCOTI, SIMONE  
; TITLE OF INVENTION: METHOD FOR IN VITRO DIAGNOSIS OF ENDOMETRIOSIS  
; FILE REFERENCE: SCH-1789  
; CURRENT APPLICATION NUMBER: US/09/961.403  
; CURRENT FILING DATE: 2001-09-25  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 2386  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-961-403-1

Query Match	97.5%;	Score	3595.5;	DB 2;	Length	2386;			
Best Local Similarity	97.2%;	Pred. No.	2.6e-301;						
Matches	629;	Conservative	1;	Mismatches	2;	Indels	15;	Gaps	1;
Qy	1	MLRGPGLLLLLAVQCLGTA	VPSTGASKSRQAQQMVQPSQSPVAVSOSKPCGYDNGKHQ	60					
Db	1	MLRGPGLLLLLAVQCLGTA	VPSTGASKSRQAQQMVQPSQSPVAVSOSKPCGYDNGKHQ	60					
Qy	61	INQWERTYLGALVCTCYGSGRGFNCE	SPKAEETCFDKYTGNTYVGTIERPKDSMI	120					
Db	61	INQWERTYLGALVCTCYGSGRGFNCE	SPKAEETCFDKYTGNTYVGTIERPKDSMI	120					
Qy	121	WDCTCAGAGRGISCTIANRCHGGOSYKIGD	TWRPHETGGYMLCEVCLNGKGEWTK	180					
Db	121	WDCTCAGAGRGISCTIANRCHGGOSYKIGD	TWRPHETGGYMLCEVCLNGKGEWTK	180					
Qy	181	PIAEKCFDHAAGTSYVVGTEWEPYQGMWVDCT	CLGEGSGRITCTSRNRCNDQDTRTSY	240					
Db	181	PIAEKCFDHAAGTSYVVGTEWEPYQGMWVDCT	CLGEGSGRITCTSRNRCNDQDTRTSY	240					
Qy	241	RIGDTSWKDNRGNLLQCICTGNRGGEWKCEHRTSVQTT	SSGSGPFTDVRAAVYQPPHP	300					
Db	241	RIGDTSWKDNRGNLLQCICTGNRGGEWKCEHRTSVQTT	SSGSGPFTDVRAAVYQPPHP	300					
Qy	301	QPPYPCHCVTDGSGVYVSGNQWLKTQGNKMLCT	CLNGVSCQETAVTQTYGNSNGEPC	360					

Db 301 QPPPYGHCVTDSGVYVSGMQLKTQGNKQMLCTCLNGVSCQETA VOTYGGNSNGPC 360  
QY 361 VLPFTYNDRT-----DSTTSNYEQDQKYSFCTDHTVLVOTRGSNGALC 405  
Db 361 VLPFTYNGRTFYSCCTEGRDGHLWCSTTSNYDQDKYSFCTDHTVLVOTRGSNGALC 420  
QY 406 HFPFLYNNHNYTDCSTSEGRDNNKWCCTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRI 465  
Db 421 HFPFLYNNHNYTDCSTSEGRDNNKWCCTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRI 480  
QY 466 GDQWDKQHDHGMHMRCTCVNGRGEWTCIAYSOLRQDCIYVDDITYNVNDTFHKKRHEEGHM 525  
Db 481 GDQWDKQHDHGMHMRCTCVNGRGEWTCIAYSOLRQDCIYVDDITYNVNDTFHKKRHEEGHM 540  
QY 526 LNCTCFGQGRGWKCDPVDQCDSETGTFYQIGDSWEKYVHGVRYQCYCGRGIGSEWHCQ 585  
Db 541 LNCTCFGQGRGWKCDPVDQCDSETGTFYQIGDSWEKYVHGVRYQCYCGRGIGSEWHCQ 600  
QY 586 PLQTPYSSSGPVEVFITETPSQNSHPIONNAPQPSHISKYILLRWRP 632  
Db 601 PLQTPYSSSGPVEVFITETPSQNSHPIONNAPQPSHISKYILLRWRP 647

RESULT 2  
US-09-016-366A-12  
; Sequence 12, Application US/09016366A  
; Patent No. 5955431  
; GENERAL INFORMATION:  
; APPLICANT: Stevens, Richard L.  
; APPLICANT: Huang, Chifu  
; TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE  
; TITLE OF INVENTION: INHIBITORS  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wolf, Greenfield & Sacks, P. C.  
; STREET: 600 Atlantic Avenue  
; CITY: Boston  
; STATE: MA  
; COUNTRY: U.S.A.  
; ZIP: 02210-2211  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/016,366A  
; FILING DATE: January 30, 1998  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/037,090  
; FILING DATE: 05-FEB-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Plumer, Elizabeth R.  
; REGISTRATION NUMBER: 36,637  
; REFERENCE/DOCKET NUMBER: B0801/7093  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-720-3500  
; TELEFAX: 617-720-2441  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2386 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-016-366A-12

Query Match 97.1%; Score 3579.5; DB 1; Length 2386;  
Best Local Similarity 96.8%; Pred. No. 6.3e-300;  
Matches 626; Conservative 1; Mismatches 5; Indels 15; Gaps 1;

QY 1 MLRGPGLLLALLVOCIGTAVPSTGASKSKRAQQMVPQSPVAVSQSKPGCYDNGKHQY 60  
Db 1 MLRGPGLLLALLVOCIGTAVPSTGASKSKRAQQMVPQSPVAVSQSKPGCYDNGKHQY 60  
QY 61 INQWERTYLGNALVCTCYGSGRGNCEKSPAEETCFDKYTGNTYRVGDTYERPKDSMI 120  
Db 61 INQWERTYLGNALVCTCYGSGRGNCEKSPAEETCFDKYTGNTYRVGDTYERPKDSMI 120  
QY 121 WDTCTCIGAGRGISCTIANRCHEGGQSYKIGDTRRRPHETGGYMLCEVCLNGKGEWTK 180  
Db 121 WDTCTCIGAGRGISCTIANRCHEGGQSYKIGDTRRRPHETGGYMLCEVCLNGKGEWTK 180  
QY 181 PLAEKCFDHAAGTSYVVGETWEKPYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 240  
Db 181 PLAEKCFDHAAGTSYVVGETWEKPYQGMWMDCTCLGEGSGRITCTSRNRCNDQDTRTSY 240  
QY 241 RIGDTWSKDNKGNLLQCICTGNRGEWKCEHRTSVQTTSSGSGPFTDVRAAVYQPPHP 300  
Db 241 RIGDTWSKDNKGNLLQCICTGNRGEWKCEHRTSVQTTSSGSGPFTDVRAAVYQPPHP 300  
QY 301 QPPPYGHCVTDSGVYVSGMQLKTQGNKQMLCTCLNGVSCQETA VOTYGGNSNGPC 360  
Db 301 QPPPYGHCVTDSGVYVSGMQLKTQGNKQMLCTCLNGVSCQETA VOTYGGNSNGPC 360  
QY 361 VLPFTYNDRT-----DSTTSNYEQDQKYSFCTDHTVLVOTRGSNGALC 405  
Db 361 VLPFTYNGRTFYSCCTEGRDGHLWCSTTSNYEQDQKYSFCTDHTVLVOTRGSNGALC 420  
QY 406 HFPFLYNNHNYTDCSTSEGRDNNKWCCTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRI 465  
Db 421 HFPFLYNNHNYTDCSTSEGRDNNKWCCTTQNYDADQKFGFCPMAAHEEICTTNEGVMYRI 480  
QY 466 GDQWDKQHDHGMHMRCTCVNGRGEWTCIAYSOLRQDCIYVDDITYNVNDTFHKKRHEEGHM 525  
Db 481 GDQWDKQHDHGMHMRCTCVNGRGEWTCIAYSOLRQDCIYVDDITYNVNDTFHKKRHEEGHM 540  
QY 526 LNCTCFGQGRGWKCDPVDQCDSETGTFYQIGDSWEKYVHGVRYQCYCGRGIGSEWHCQ 585  
Db 541 LNCTCFGQGRGWKCDPVDQCDSETGTFYQIGDSWEKYVHGVRYQCYCGRGIGSEWHCQ 600  
QY 586 PLQTPYSSSGPVEVFITETPSQNSHPIONNAPQPSHISKYILLRWRP 632  
Db 601 PLQTPYSSSGPVEVFITETPSQNSHPIONNAPQPSHISKYILLRWRP 647

RESULT 3  
US-08-551-356-2  
; Sequence 2, Application US/08551356  
; Patent No. 5830700  
; GENERAL INFORMATION:  
; APPLICANT: Irani, Meher  
; TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ZymoGenetics, Inc.  
; STREET: 4225 Roosevelt Way, N.E.  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/551,356  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/349,762  
; FILING DATE:  
; APPLICATION NUMBER: US/07/998,271

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 10, 2006, 19:13:10 ; Search time 5.93558 Seconds  
(without alignments)  
740.245 Million cell updates/sec

Title: US-09-581-651D-41

Perfect score: 54

Sequence: 1 VSIPPNLGY 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 21.\*

1: Geneseqp1980s.\*

2: Geneseqp1990s.\*

3: Geneseqp2000s.\*

4: Geneseqp2001s.\*

5: Geneseqp2002s.\*

6: Geneseqp2003as.\*

7: Geneseqp2003bs.\*

8: Geneseqp2004s.\*

9: Geneseqp2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	100.0	20	2 AAY28902	Aay28902 MSF 1-alp
2	54	100.0	642	8 ADQ39403	Adq39403 Human myo
3	54	100.0	642	8 ADR67316	Adr67316 Human bla
4	54	100.0	642	8 ADS17489	Adsl17489 Amino aci
5	54	100.0	642	8 ADR97658	Adr97658 Human fib
6	54	100.0	642	9 ADW44480	Adw44480 Human fib
7	54	100.0	642	9 ADY55704	Ady55704 Human fib
8	54	100.0	642	9 ADY32458	Ady32458 Human fib
9	54	100.0	642	9 AEB56258	Aeb56258 Event seq
10	54	100.0	642	9 AEB78062	Aeb78062 Human fib
11	54	100.0	657	8 ADQ39409	Adq39409 Human myo
12	54	100.0	657	9 ADZ26739	Adz26739 Human fib
13	54	100.0	660	2 AAY28901	Aay28901 Human mig
14	40	74.1	291	2 AAR65443	Aar65443 Variant 1
15	38	70.4	145	4 ABB71381	Abb71381 Drosophil
16	38	70.4	279	7 ADF29037	Adf29037 M. globos
17	38	70.4	291	5 AAO15869	AAo15869 Talaromyc
18	38	70.4	291	5 AAO19509	AAo19509 T thermop
19	38	70.4	351	7 ABO79631	ABo79631 Pseudom
20	38	70.4	500	4 AAU02826	AAu02826 Taxus cus
21	38	70.4	502	9 ADW72676	Adw72676 Taxus cus
22	37	68.5	40	6 ADA56768	Ada56768 Human sec
23	37	68.5	40	6 ADA40618	Ada40618 Human sec
24	37	68.5	40	7 ADB91474	ADB91474 Human sec

Query Match 100.0% Score 54; DB 2; Length 20;

25	37	68.5	40	8	ADN41076	Adn41076 Novel hum
26	37	68.5	41	2	AAY02711	Aay02711 Human sec
27	37	68.5	41	7	ADA07390	Ada07390 Human sec
28	37	68.5	122	4	AAO06277	Aao06277 Human pol
29	37	68.5	362	7	ABM88358	Abm88358 Rice abio
30	37	68.5	485	4	RAU02835	Rau02835 Taxus cus
31	37	68.5	485	9	ADM72680	Adm72680 Taxus cus
32	37	68.5	515	8	ADN47628	Adn47628 Thermococ
33	37	68.5	3482	9	ADV24897	Adv24897 Anabaena
34	36	66.7	98	4	ABB16357	Abb16357 Human ner
35	36	66.7	150	4	ABG05775	Abg05775 Novel hum
36	36	66.7	154	3	RAB54331	Rab54331 Human pan
37	36	66.7	216	7	ADE61648	Ade61648 Rat Prote
38	36	66.7	243	5	ABP59149	Abp59149 Ribosomal
39	36	66.7	269	4	AAB92540	Aab92540 Human pro
40	36	66.7	281	5	ABB97335	Abb97335 Novel hum
41	36	66.7	291	2	AAR65437	Aar65437 Variant 1
42	36	66.7	294	8	ADR09348	Adr09348 Human pro
43	36	66.7	359	7	ABM88720	Abm88720 Rice abio
44	36	66.7	512	4	AAU02834	Aau02834 Taxus cus
45	36	66.7	578	4	ABG02192	Abg02192 Novel hum

#### ALIGNMENTS

##### RESULT 1

AAV28902

ID AAV28902 standard; peptide; 20 AA.

XX AAV28902;

XX AAV28902;

DT 21-SEP-1999 (first entry)

XX MSF 1-alpha peptide epitope.

DE MSF 1-alpha peptide epitope.

XX Migration stimulatory factor; MSF; cell migration; modulation; human;

KW wound healing; scarring; MSF1-alpha; epitope; fibronectin.

XX Synthetic.

OS Homo sapiens.

XX WO9931233-A1.

PN 24-JUN-1999.

XX 15-DEC-1998; 98WO-GB003766.

PF 16-DEC-1997; 97GB-00026539.

XX (UYDU-) UNIV DUNDEE.

XX Schor SL, Schor AM;

XX WPI; 1999-430039/36.

XX Proteins with cell migration stimulatory activity used in treating wound

XX and preventing scarring.

XX Claim 18; Page 57; 86pp; English.

XX The invention provides a human migration stimulatory factor (MSF)

CC protein. Host cells containing a replicable vector comprising the MSF

CC encoding nucleic acid can be used for the recombinant production of the

CC protein. The polypeptide can be used for modulating cell migration,

CC healing a wound and for preventing scarring. Sequences AAY28902-906

CC represent peptide epitopes of MSF against which monoclonal antibodies

CC that are specific to MSF and do not cross-react with fibronectin are

XX raised

XX Sequence 20 AA;





Db 633 VSIPPRNLGY 642

RESULT 4  
ADS17489  
ID ADS17489 standard; protein; 642 AA.  
XX AC ADS17489;  
XX DT 02-DEC-2004 (first entry)  
XX DE Amino acid sequence of human fibronectin 1.  
XX KW cell state; time-lapse profile; protein-protein interaction;  
KW drug screening; cancer; infectious disease; allergy; hypertension;  
KW hyperlipaemia; diabetes; cardiac disease; cerebral infarction; dementia;  
KW obesity; arteriosclerosis; infertility; mental disease; nervous disease;  
KW cataract; progeria; hypersensitivity; ultraviolet radiation; human;  
KW fibronectin 1; actin acting substance; transfection array.  
XX OS Homo sapiens.  
XX WO2004079007-A2.  
XX PN 16-SEP-2004.  
XX PD 03-MAR-2004; 2004WO-JP002694.  
XX PF 04-MAR-2003; 2003JP-00057870.  
XX PR (NAAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.  
XX PA Miyake M, Yoshikawa T, Uchimura E, Miyake J;  
XX WPI; 2004-662438/64.  
XX DR N-PSDB; ADS17488.  
XX PS Presenting a state of a cell, useful for diagnosing and treating a  
PT disease, e.g. cancer, infectious disease, allergy, diabetes, dementia,  
PT obesity, infertility, or cataract, comprises obtaining a time-lapse  
PT profile of the cell.  
XX Disclosure; SEQ ID NO 2; 532pp; English.

XX The specification describes a method and system for accurately presenting  
CC a state of a cell. The method comprises obtaining a time-lapse profile of  
CC the cell by time-lapse monitoring of a gene state associated with at  
CC least one gene derived from the cell, and presenting the time-lapse  
CC profile. The gene comprises a transcription control sequence, and the  
CC gene state includes expression of the gene. The method and system are  
CC useful presenting a state of a cell. The method can allow the elucidation  
CC of key protein-protein interactions suitable for targeting by drug  
CC screening protocols. The method is useful for diagnosing or treating a  
CC disease, e.g. cancer, infectious disease due to viruses or bacteria,  
CC allergy, hypertension, hyperlipaemia, diabetes, cardiac disease, cerebral  
CC infarction, dementia, obesity, arteriosclerosis, infertility, mental and  
CC nervous diseases, cataract, progeria, or hypersensitivity to ultraviolet  
CC radiation. The present sequence represents human fibronectin 1. Bovine  
CC fibronectin was used as a candidate for an actin acting substance. The  
CC actin acting substance was used with transfection reagents and amplified  
CC plasmid DNA in assays using transfection arrays, in the course of the  
CC invention.

SQ Sequence 642 AA;  
Query Match 100.0%; Score 54; DB 8; Length 642;  
Best Local Similarity 100.0%; Pred. No. 0.51;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VSIPPRNLGY 10  
| | | | | | | |  
Db 633 VSIPPRNLGY 642

RESULT 5  
ADR97658  
ID ADR97658 standard; protein; 642 AA.  
XX AC ADR97658;  
XX DT 02-DEC-2004 (first entry)  
XX DE Human fibronectin 1 protein, an actin acting substance SeqID 2.  
XX KW human; transfection efficiency; actin acting substance;  
KW extracellular matrix; fibronectin 1; gene introduction reagent.  
XX OS Homo sapiens.  
XX WO2004079332-A2.  
XX PN 16-SEP-2004.  
XX PD 03-MAR-2004; 2004WO-JP002696.  
XX PF 04-MAR-2003; 2003JP-00057869.  
XX PR (NAAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.  
XX PA Miyake M, Yoshikawa T, Uchimura E, Miyake J;  
XX WPI; 2004-677173/66.  
XX DR N-PSDB; ADR97657.  
XX PS Composition comprising an actin acting substance or an actin acting  
PT substance and a target substance, useful for increasing the efficiency of  
PT introducing a target substance into a cell and in cell biology or genetic  
PT engineering.  
XX Claim 4; SEQ ID NO 2; 347pp; English.

XX This invention relates to a novel composition and method for increasing  
CC the efficiency of introducing a target substance into a cell.  
CC Specifically, it refers to the introduction of DNA (e.g. the gene of  
CC interest to be transfected), polypeptides, sugars or complexes thereof  
CC into a cell, and comprises an actin acting substance. The present  
CC invention describes the actin acting substance as an extracellular matrix  
CC protein, a variant or fragment thereof selected from fibronectin, laminin  
CC or vitronectin. The composition further comprises a gene introduction  
CC reagent selected from cationic polymers, cationic lipids, and calcium  
CC phosphate, as well as a gold colloid particle that is contacted with the  
CC cell. As such, the composition, kit, device or method is useful for  
CC increasing the efficiency of introducing a target substance into a cell  
CC and thus is useful in the fields of cell biology, genetic engineering and  
CC molecular biology. This polypeptide sequence is the human fibronectin  
CC protein sequence of the invention.

SQ Sequence 642 AA;  
Query Match 100.0%; Score 54; DB 8; Length 642;  
Best Local Similarity 100.0%; Pred. No. 0.51;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VSIPPRNLGY 10  
| | | | | | | |  
Db 633 VSIPPRNLGY 642

RESULT 6  
ADW44480  
ID ADW44480 standard; protein; 642 AA.  
XX AC ADW44480;  
XX DT 24-MAR-2005 (first entry)  
XX

```
DE Human fibronectin 1.
XX cell transduction; nerves; cell adhesion; fibronectin 1.
XX Homo sapiens.
XX WO2005001090-A1.
XX 06-JAN-2005.
XX 25-JUN-2004; 2004WO-JP009568.
XX 26-JUN-2003; 2003JP-00183630.
XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX Miyake M, Uchimura E, Yoshikawa T, Miyake J;
XX WPI; 2005-091515/10.
XX DR N-PSDB; ADW44479.
XX Composition useful for improving transduction efficiency of nucleic acid
XX into cell, comprises cell adhesion molecule and gene-transfer reagent.
XX Disclosure; SEQ ID NO 24; 446pp; Japanese.
XX This invention describes a novel composition for improving transduction
XX efficiency of a nucleic acid into a cell from nervous tissue which
XX comprises a cell adhesion molecule and a gene-transfer reagent. The
XX invention also describes a device, kit and novel method for improving
XX nucleic acid transduction efficiency of a cell on a solid phase. The cell
XX adhesion molecule contains an extracellular matrix chosen from collagen,
XX laminin and fibronectin. The gene transfer reagent comprises a cationic
XX polymer, cationic liquid, polyamine type reagent, polyamine type reagent
XX or calcium phosphate. This sequence represents human fibronectin 1, which
XX is used in the composition of the invention.
XX SQ Sequence 642 AA;
    Query Match 100.0%; Score 54; DB 9; Length 642;
    Best Local Similarity 100.0%; Pred. No. 0.51;
    Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VSIPPRNLGY 10
Db 633 VSIPPRNLGY 642
RESULT 7
ADY55704
ID ADY55704 standard; protein; 642 AA.
XX ADY55704;
XX 19-MAY-2005 (first entry)
XX Human fibronectin 1 protein.
XX biocomputer; drug screening; diagnosis; food; cosmetics; agriculture;
XX analysis; fibronectin.
XX Homo sapiens.
XX WO2005021744-A1.
XX 10-MAR-2005.
XX 25-JUN-2004; 2004WO-JP009404.
XX 25-JUN-2003; 2003JP-00181915.
XX 07-AUG-2003; 2003JP-00289469.
XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA
```

---

```
XX Miyake M, Yoshikawa T, Miyake J;
XX WPI; 2005-262218/27.
XX DR N-PSDB; ADY55703.
XX Performing data production and presenting cell information under a
XX similar environment as digital cells, useful e.g. in drug development,
XX comprises profiling the actual status of cells.
XX Example 14; SEQ ID NO 2; 517pp; Japanese.
XX The invention relates to a novel method for producing profile data that
XX relates to cell information in an identical environment. The method
XX comprises placing several cells on a support in the same environment, and
XX monitoring the biological factors on or in the cells or their aggregate
XX chronologically to generate profile data of the cells. The methods and
XX systems are useful for performing data production by profiling actual
XX status of cells, and for presenting time-lapse and/or real-time cell
XX information at a cellular level under a similar environment as well as
XX digital cells in the form of a library of databases with cell
XX information, which are particularly useful in drug development, disease
XX diagnosis and management, as well as in fields including food technology,
XX cosmetics, agriculture, environmental sciences, in silico computational
XX biological studies, cell analysis and biotechnology. The methods and
XX systems are capable of providing information with controllability,
XX specificity and most importantly directly even in complex systems. This
XX sequence represents a human fibronectin 1 protein used in the method for
XX producing a digital cell profile of the invention.
XX SQ Sequence 642 AA;
    Query Match 100.0%; Score 54; DB 9; Length 642;
    Best Local Similarity 100.0%; Pred. No. 0.51;
    Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VSIPPRNLGY 10
Db 633 VSIPPRNLGY 642
RESULT 8
ADY32458
ID ADY32458 standard; protein; 642 AA.
XX ADY32458;
XX 19-MAY-2005 (first entry)
XX Human fibronectin protein.
XX biochip; fluorescence; fibronectin.
XX Homo sapiens.
XX JP2005052050-A.
XX 03-MAR-2005.
XX 01-AUG-2003; 2003JP-00285471.
XX 01-AUG-2003; 2003JP-00285471.
XX (DOKU-) DOKURITSU GYOSEI HOJIN SANGYO GIJUTSU SO.
XX WPI; 2005-199806/21.
XX DR N-PSDB; ADY32457.
XX Disc-like biochip useful for printing biomolecules such as DNA, comprises
XX cell fixed to disc-like solid phase support body.
XX Claim 16; SEQ ID NO 2; 43pp; Japanese.
XX
```

CC The invention relates to a novel disc-like biochip, comprising a cell  
 CC fixed to a disc-like solid phase support body. The invention further  
 CC comprises: a reader of the disc-like biochip in order to read the spot on  
 CC the chip and a detection unit to detect the signal produced from the  
 CC spot. The disc-like biochip is useful for printing biomolecules such as  
 CC DNA at high speed. The reader of the biochip performs fluorescent reading  
 CC in a cheap and simple manner. This sequence represents a human  
 CC fibronectin protein used in the novel biochip detection device of the  
 CC invention.

XX SQ Sequence 642 AA;

Query Match 100.0%; Score 54; DB 9; Length 642;

Best Local Similarity 100.0%; Pred. No. 0.51; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSIPPRNLGY 10  
 DB 633 VSIPPRNLGY 642

RESULT 9

AEBS6258  
 ID AEB56258 standard; protein; 642 AA.

XX AC AEB56258;

XX DT 20-OCT-2005 (first entry)

XX DE Event sequencer related protein SEQ ID NO 2.

XX KW analyte detection.

XX OS Unidentified.

XX PN W02005073890-A1.

XX PD 11-AUG-2005.

XX PF 27-JAN-2005; 2005WO-JP001151.

XX PR 30-JAN-2004; 2004JP-00024923.

XX PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.

XX PI Miyake M, Yoshikawa T, Miyake J;

XX DR WPI; 2005-571425/58.

XX DR N-PSDB; AEB56257.

XX PT Event sequencer extracts portion of timer series data with peculiar  
 PT behavior associated with index of system, as event timing, to generate  
 PT event descriptor described by event timing.

XX PS Disclosure; SEQ ID NO 2; 281pp; Japanese.

XX CC The invention relates to an event sequencer that acquires time-series  
 CC data on an index derived from a system and provides a peculiar behavior  
 CC associated with the index. A portion of time series data having the  
 CC peculiar behavior is extracted as an event timing and an event descriptor  
 CC described by the event timing is generated. The sequencer is useful for  
 CC analysis of the state of a system. The sequencer performs meaningful  
 CC analysis of a system state using specific index effectively. The present  
 CC sequence represents the amino acid sequence of an event sequencer related  
 CC protein.

XX SQ Sequence 642 AA;

Query Match 100.0%; Score 54; DB 9; Length 642;

Best Local Similarity 100.0%; Pred. No. 0.51; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSIPPRNLGY 10

DB 633 VSIPPRNLGY 642

RESULT 10  
 AEB78062

ID AEB78062 standard; protein; 642 AA.

XX AC AEB78062;

XX DT 20-OCT-2005 (first entry)

XX DE Human fibronectin 1.

XX KW Transfection; fibronectin; cell adhesion; gene transfer.

XX OS Homo sapiens.

XX PN W02005073385-A1.

XX PD 11-AUG-2005.

XX PF 27-JAN-2005; 2005WO-JP001148.

XX PR 29-JAN-2004; 2004JP-00022315.

XX PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.

XX PI Miyake M, Yoshikawa T, Uchimura E, Miyake J;

XX DR WPI; 2005-555702/56.

XX DR N-PSDB; AEB78061.

XX PT Composition for elevating transfer efficiency of target substance into  
 PT cells, comprises cell adhesion-related factor.

XX PS Disclosure; SEQ ID NO 2; 101pp; Japanese.

XX CC The invention relates to a composition (I) for elevating the transfer  
 CC efficiency of a target substance into cells, comprises a cell adhesion-  
 CC related factor. Also included are a kit (K1) for elevating gene transfer  
 CC efficiency (comprising cell-adhesion related factor, and a gene transfer  
 CC reagent), a composition (II) for introducing a target substance into a  
 CC cell (comprising a target substance, and a cell-adhesion related factor),  
 CC a device (III) for elevating transfer efficiency of target substance into  
 CC cell (comprising a target substance and a cell-adhesion related factor,  
 CC where the cell-adhesion related factor was immobilized to a support) and  
 CC elevating (M1) the transfer efficiency of a target substance into cells  
 CC (involving providing a target substance and a cell-adhesion related  
 CC factor, and contacting the target substance and the cell-adhesion related  
 CC factor with the cell). The cell-adhesion related factor contains an  
 CC interaction substance capable of interacting with cell adhesion molecule  
 CC (e.g. the extracellular matrix). The cell adhesion molecule contains  
 CC integrin receptor or RGD motif. The interaction substance carries out an  
 CC antigen antibody reaction to partner of the cell adhesion molecule. The  
 CC interaction substance is an antibody or its derivative(s). The  
 CC interaction substance contains an antibody chosen from anti-CD49a  
 CC antibody, anti-CD49b antibody, anti-CD49c antibody, anti-CD49e antibody  
 CC and anti-CD49f antibody. The integrin receptor is chosen from CD49a,  
 CC CD49b, CD49c, CD49d, CD49e, CD49f and CD29. The integrin receptor interacts with  
 CC CD49c, CD49d, CD49e, CD49f and CD29. The integrin receptor interacts with  
 CC molecule chosen from collagen, fibronectin, vitronectin and laminin. (I),  
 CC (K1), (II) or (M1) are useful for elevating the transfer efficiency of a  
 CC target substance into cells, where the target substance is a genetic  
 CC material (nucleic acid molecule), preferably DNA, and the cell is at  
 CC least one cell chosen from stem cell and differentiation cell. (II) is  
 CC useful for introducing a target substance into a cell, where the target  
 CC substance is chosen from DNA, RNA, polypeptide, saccharide, and their  
 CC complex, preferably DNA. (I) enables to improve the transfer efficiency  
 CC under any conditions (solid or liquid phase) in the case of transferring  
 CC a target substance that can be hardly transferred into cells. The present  
 CC sequence represents human fibronectin 1. Note: the sequence data for this  
 CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 642 AA;

Query Match 100.0%; Score 54; DB 9; Length 642;  
Best Local Similarity 100.0%; Pred. No. 0.51;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSIPPRNLGY 10  
Db 633 VSIPPRNLGY 642  
|||||

RESULT 11  
ADQ39409  
ID ADQ39409 standard; protein; 657 AA.  
XX  
AC ADQ39409;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Human myocardial infarction-associated gene derived protein, SEQ ID 1072.  
XX  
KW Myocardial infarction; detection; single nucleotide polymorphism; SNP;  
KW cardiant; gene therapy; human.  
XX  
OS Homo sapiens.  
XX  
PN WO2004058052-A2.  
XX  
PD 15-JUL-2004.  
XX  
PF 22-DEC-2003; 2003WO-US040978.  
XX  
PR 20-DEC-2002; 2002US-0434778P.  
PR 10-MAR-2003; 2003US-0453135P.  
PR 30-APR-2003; 2003US-0466412P.  
PR 23-SEP-2003; 2003US-0504955P.  
XX  
PA (APPL-) APPLERA CORP.  
XX  
PI Cargill M, Devlin JJ, Takoubova O;  
XX  
DR WPI; 2004-533949/51.  
DR N-PSDB; ADQ38581.  
XX

Identifying an individual who has an altered risk for developing  
PT myocardial infarction by detecting a single nucleotide polymorphism in  
PT the individual's nucleic acids.  
XX  
PS Claim 10; SEQ ID NO 1072; 145pp; English.  
XX

The invention relates to a novel method for identifying an individual who  
CC has an altered risk for developing myocardial infarction. The method  
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of  
CC the nucleotide sequences given in the specification in the individual's  
CC nucleic acids, where the presence of the SNP is correlated with an  
CC altered risk for myocardial infarction in the individual. The invention  
CC further comprises: an isolated nucleic acid molecule comprising at least  
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in  
CC the specification or its complement and encoding any one of the amino  
CC acid sequences given in the specification; an isolated polypeptide  
CC comprising an amino acid sequence given in the specification; an antibody  
CC that specifically binds to the polypeptide or its antigen-binding  
CC fragment; an amplified polynucleotide containing an SNP given in the  
CC specification and which is between about 16 and 1000 nucleotides in  
CC length; a kit for detecting an SNP in a nucleic acid, comprising the  
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a  
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a  
CC method for identifying an agent useful in treating or preventing  
CC myocardial infarction. The novel detection method has cardiant activity.  
CC The nucleic acids of the invention may be used in gene therapy. The

method is useful in identifying an individual who has an increased or  
CC decreased risk for developing myocardial infarction and for preparing a  
CC composition for treating or preventing myocardial infarction. This  
CC sequence represents the protein of a human myocardial infarction-  
CC associated gene containing one or more SNP's of the invention. Note: This  
CC sequence was not shown in the specification. The sequence has come from  
CC an electronic sequence listing downloaded from the WIPO website.

XX  
SQ Sequence 657 AA;

Query Match 100.0%; Score 54; DB 8; Length 657;  
Best Local Similarity 100.0%; Pred. No. 0.53;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSIPPRNLGY 10  
Db 648 VSIPPRNLGY 657  
|||||

RESULT 12  
AD226739  
ID AD226739 standard; protein; 657 AA.  
XX  
AC AD226739;  
XX  
DT 16-JUN-2005 (first entry)  
XX  
DE Human fibronectin.  
XX  
KW cell culture; stem cell; fibronectin.  
XX  
OS Homo sapiens.  
XX  
PN WO2005030999-A1.  
XX  
PD 07-APR-2005.  
XX  
PF 24-SEP-2004; 2004WO-US031524.  
XX  
PR 25-SEP-2003; 2003US-0506221P.  
PR 08-OCT-2003; 2003US-0509594P.  
XX  
PA (DAND ) DANA FARBER CANCER INST INC.  
XX  
PI Ritz J, Wu CJ;  
XX  
DR WPI; 2005-273394/28.  
DR N-PSDB; AD226738.  
XX

Detecting lineage-specific cells in a biological sample, useful for  
PT determining the clinical outcome of a progenitor cell transfer in a  
PT subject, comprises identifying lineage-specific mRNA in the sample.  
XX  
PS Disclosure; SEQ ID NO 365; 393pp; English.  
XX

The invention relates to a method of detecting lineage-specific cells in  
CC a biological sample which comprises identifying lineage-specific mRNA in  
CC the sample. The methods are useful for determining the clinical outcome  
CC of a progenitor cell transfer in a subject, and for identifying or  
CC quantifying lineage-specific cells. The present sequence represents the  
CC amino acid sequence of a human protein used to identify lineage-specific  
CC cells.

XX  
SQ Sequence 657 AA;

Query Match 100.0%; Score 54; DB 9; Length 657;  
Best Local Similarity 100.0%; Pred. No. 0.53;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSIPPRNLGY 10  
Db 648 VSIPPRNLGY 657  
|||||

```

RESULT 13
AA28901
ID AAY28901 standard; protein; 660 AA.
XX
XX AAY28901;
AC
XX
XX 21-SEP-1999 (first entry)
DT
XX
XX Human migration stimulating factor (MSF) 1-alpha protein.
DE
XX
XX Migration stimulatory factor; MSF; cell migration; modulation; human;
KW wound healing; scarring; MSF1-alpha.
KW
XX
XX Homo sapiens.
OS
XX
XX WO9931233-A1.
PN
XX
XX 24-JUN-1999.
PD
XX
XX 15-DEC-1998; 98WO-GB003766.
PF
XX
XX 16-DEC-1997; 97GB-00026539.
PR
XX
XX (UYDU-) UNIV DUNDEE.
PA
XX
XX Schor SL, Schor AM;
PI
XX
XX WPI; 1999-430039/36.
DR
XX
XX N-PSDB; AAX81299.
DR
XX
XX
XX
XX Proteins with cell migration stimulatory activity used in treating wound
PT and preventing scarring.
FT
XX
XX Claim 1; Page 53; 86pp; English.
PS
XX
XX The invention provides a human migration stimulatory factor (MSF)
CC protein. Host cells containing a replicable vector comprising the MSF
CC encoding nucleic acid can be used for the recombinant production of the
CC protein. The polypeptide can be used for modulating cell migration,
CC healing a wound and for preventing scarring. The present sequence
CC represents the human MSF1-alpha protein
XX
XX Sequence 660 AA;
SQ
Query Match 100.0%; Score 54; DB 2; Length 660;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VSIPPRNLGY 10
DB 651 VSIPPRNLGY 660
|||||
|

RESULT 14
AA65443
ID AAR65443 standard; protein; 291 AA.
XX
XX AAR65443;
AC
XX
XX 25-MAR-2003 (revised)
DT
DT 17-JUL-1995 (first entry)
XX
XX Variant lipase D96W+E210N, used in detergent compositions.
DE
XX
XX lipase; variant; catalytic triad; lipid contact zone; active Serine;
KW wash performance; detergent; dishwashing; softening composition.
KW
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH 23..291
FT Protein /label= mature_lipase

```

---

```

FT Region 43..47
FT /note= "lipid contact zone"
FT Region 48..50
FT /note= "lipid contact zone"
FT Region 78..84
FT /note= "lipid contact zone"
FT Region 103..120
FT /note= "lipid contact zone"
FT Misc-difference 118
FT /label= D96W
FT Region 132..138
FT /note= "lipid contact zone"
FT Region 166..169
FT /note= "lipid contact zone"
FT Region 194..198
FT /note= "lipid contact zone"
FT Region 221..235
FT /note= "lipid contact zone"
FT Misc-difference 232
FT /label= E210N
FT Region 270..291
FT /note= "lipid contact zone"
FT
XX WO9425577-A1.
PN
XX
XX 10-NOV-1994.
PD
XX
XX 22-APR-1994; 94WO-DK000162.
PF
XX
XX 23-APR-1993; 93DK-00000466.
PR
XX
XX (NOVO ) NOVO-NORDISK AS.
PA
XX
XX Svendsen A, Patkar SA, Gormsen E, Clausen IG;
PI WPI; 1994-358256/44.
XX
XX New lipase variant with non-aromatic amino acid replaced - in the lipid
PT contact zone, and related DNA, vectors and transformed cells, useful in
PT detergent compens. etc.
XX
XX Claim 6; Page 32-34; 44pp; English.
XX
XX This polypeptide is a variant of the Humicola lanuginosa lipase shown in
CC AAR65394. This variant contains a double mutation, where at position 96,
CC the wild type aspartic acid is substituted with tryptophan and at
CC position 210, the wild type glutamic acid is substituted with asparagine.
CC Variants are constructed in which a non-aromatic amino acid in the lipid
CC contact zone is replaced with an aromatic amino acid (see also AAR65396-
CC 442 and AAR65444-449). The parent lipase has a trypsin-like catalytic
CC triad including an active serine in a hydrophobic, elongated binding
CC pocket in the lipid contact zone (located in the part of the lipase
CC structure contg. the active Ser and involved in interaction with the
CC substrate at or during hydrolysis). Some variants were constructed in
CC which one or more amino acid residues were replaced in specific positions.
CC The variants are useful as additives for detergent, dishwashing and
CC softening compens. They provide better wash performance (with
CC improvement factors as high as 4) than wild type enzymes. (Updated on 25-
CC MAR-2003 to correct PN field.)
XX
XX Sequence 291 AA;
SQ
Query Match 74.1%; Score 40; DB 2; Length 291;
Best Local Similarity 75.0%; Pred. No. 78;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 IPPRNLGY 10
DB 228 LPPRNFY 235
|||||
|

RESULT 15
ABB71361

```

ID ABB71381 standard; protein; 145 AA.  
XX  
AC ABB71381;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 40935.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
XX pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
XX  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR WPI; 2001-656860/75.  
XX  
DR N-PSDB; ABL15484.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions.  
XX  
PS Disclosure; SEQ ID NO 40935; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
CC ABB72072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 145 AA;  
  
Query Match 70.4%; Score 38; DB 4; Length 145;  
Best Local Similarity 77.8%; Pred. No. 86;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 VSIPPRNLG 9  
DB 81 VSIPPPNMG 89  
  
Search completed: May 10, 2006, 19:20:06  
Job time : 10.4356 secs